Digitized by Arya Samaj Foundation Chennai and eGangotri

Digitized by Arya Samaj Foundation Chennai and eGangotri

111678

गुरुकुल कांगड़ी विश्वविद्यालय, हरिद्वार
पुस्तकालय

R
530.5
AM3P
V.134
No-4-6
67032

विषय संख्या
पुस्तक संख्या
आगत पंजिका संख्या

पुस्तक पर किसी प्रकार का निशान लगाना वर्जित है। कृपया १५ दिन से अधिक समय तक पुस्तक अपने पास न रखें।

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

67032 RBL

THE PHYSICAL REVIEW

VOLUME 134 Second Series NUMBER 4A–6

18 MAY 1964

एक प्रति निःशुल्क वितरण हेतु

Published for

THE AMERICAN PHYSICAL SOCIETY

by the

AMERICAN INSTITUTE OF PHYSICS

LANCASTER, PA., AND NEW YORK, N. Y.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

R
530.5
Am3P
V.134
No.4A-6A

THE PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893.

Published by the
American Institute of Physics
for
The American Physical Society

Officers

President
R. F. BACHER
California Institute of Technology

Vice-President
FELIX BLOCH
Stanford University

Vice-President Elect
J. A. WHEELER
Princeton University

Secretary
K. K. DARROW
Bell Telephone Laboratories, Inc.
(*retired*)

Treasurer
S. L. QUIMBY
Columbia University

Deputy Secretary
W. W. HAVENS, JR.
Columbia University

Regional Secretaries:

Central States
R. G. SACHS
University of Wisconsin

Western States
WARD WHALING
California Institute of Technology

Managing Editor
S. A. GOUDSMIT
Brookhaven National Laboratory

Office of the Society
Columbia University
New York, New York 10027

Board of Editors

Editor
S. PASTERNACK

Assistant Editor
A. HERSCHMAN

Associate Editors

1962–1964	*1963–1965*	*1964–1966*
K. R. ATKINS	R. K. ADAIR	D. A. BROMLEY
WALTER KOHN	MURRAY GELL-MANN	R. H. DALITZ
M. D. KRUSKAL	EDWARD GERJUOY	F. J. DYSON
A. V. PHELPS	A. L. SCHAWLOW	STANLEY GESCHWIND
W. W. TYLER	J. R. SCHRIEFFER	M. S. GREEN
JOSEPH WENESER	L. I. SCHIFF	P. J. PRICE

Manuscripts for publication should be submitted to S. PASTERNACK, Editor, *The Physical Review,* Brookhaven National Laboratory, Upton, New York 11973. The authors' institutions are requested to pay a publication charge of $60 per page which, if honored, entitles them to 100 free reprints. An additional $10 will be requested for prepublication of the abstract in *Physics Abstracts.* Instructions will be sent with galley proofs.

Proof and all correspondence concerning papers in the process of publication should be addressed to the Supervisor, Editorial Department, American Institute of Physics, 335 East 45 Street, New York, New York 10017. Reference must be made to title, author, journal, and scheduled date of publication.

Subscription Price

	U.S.A., Canada, and Mexico	Elsewhere
Members of The American Physical Society	$15.00	$17.00
Nonmembers	40.00	43.00

Back Numbers

Yearly back-number rate when complete year is available:

Prior to 1929, and 1942 to 1946, inclusive	**$30.00**
1929 to 1963 except 1942 to 1946, inclusive	**$45.00**
1964 and thereafter	**$50.00**

Single copies: $3.00 each. *Index issue* (1964 and thereafter): $2.00 each.

General index, 1921–1950: Paper binding, **$10** (Members); **$14** (Nonmembers)
Cloth binding, **$12** (Members); **$16** (Nonmembers)

Author index, 1956–1960: Paper binding, **$ 2** (Members); **$ 3** (Nonmembers)

Subscriptions, renewals, orders for back numbers, and changes of address should be addressed to the Circulation Department, American Institute of Physics, 335 East 45 Street, New York, New York 10017. For change of address, allow at least six weeks' advance notice; send both old and new addresses and, if possible, please include an address stencil imprint from the mailing envelope of a recent issue.

Permission is granted to quote from this journal in scientific articles, with the customary acknowledgment of the source. Reproduction of figures and tables is permitted provided that the author's consent is obtained and that the editorial office is notified. Reproduction of whole articles for commercial purposes requires written permission of both Editor and author.

The Physical Review is published weekly at Prince and Lemon Streets, Lancaster, Pennsylvania 17604.

Second-class postage paid at Lancaster, Pa.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CONTENTS

THE PHYSICAL REVIEW SECOND SERIES, VOLUME 134, NO. 4A 18 MAY 1964

Resonant Frequencies of an Electromagnetic Cavity in an Accelerated System of Reference . C. V. Heer A799
Theory of the Electronic Thermal Conductivity of Superconductors with Strong Electron-Phonon Coupling . . . . . Vinay Ambegaokar and Ludwig Tewordt A805
Dissipation in Quantum Mechanics. Two-Level System. II . . . . . I. R. Senitzky A816
Cluster Expansion in the Heitler–London Approach to Many-Electron Problems . . . . . Tadashi Arai A824
Thomas-Fermi Theory of the Atom as a Solution of the Density-Matrix Hierarchy . Nandor L. Balazs A841
Theory of Core-Electron Contributions to Hyperfine Interactions . . . . . G. D. Gaspari, Wei-Mei Shyu, and T. P. Das A852
Millimeter-Wave Molecular-Beam Spectroscopy: Alkali Chlorides . Paul L. Clouser and Walter Gordy A863
Precision Evaluation of the High-Frequency Limit of the Continuous X-Ray Spectrum Using a Gas Target X-Ray Tube . . . . . Jon J. Spijkerman and J. A. Bearden A871
Inelastic Scattering of Electrons by the Hydrogen Molecule Ion . . . . . James M. Peek A877
$\Delta S = \pm 1$ Magnetic Multipole Radiative Transitions . . . . . Masataka Mizushima A883
Absolute Electron Excitation Cross Sections of Helium . . . . . Robert M. St. John, Frank L. Miller, and Chun C. Lin A888
K-Shell Photoelectric Cross Sections from 200 keV to 2 MeV . . . . . R. H. Pratt, Richard D. Levee, Robert L. Pexton, and Walter Aron A898
Polarization Correlations in Atomic Photoeffect . . . . . R. H. Pratt, Richard D. Levee, Robert L. Pexton, and Walter Aron A916
Effect of Correlation on the Ferromagnetism of Transition Metals . . . . . Martin C. Gutzwiller A923
Mechanism of the Ferrimagnetic to Antiferromagnetic Transition in $Mn_{2-x}Cr_xSb$ . . . . . S. S. Jarrett A942
Magnetic Susceptibilities of Transition Elements in Host Crystals. II. Ni^{2+} in ZnO and CdS . . . . . William H. Brumage and Chun C. Lin A950
Electronic Specific Heats of Dilute Solid Solutions . . . . . H. Jones A958
Motional Narrowing of Nuclear Magnetic-Resonance Lines in Manganese-Doped Lithium Fluoride . . . . . T. G. Stoebe, T. O. Ogurtani, and R. A. Huggins A963
Localized Mode in an Anharmonic Crystal . . . . . William M. Visscher A965
Energy Bands for the Iron Transition Series . . . . . L. F. Mattheiss A970
Fermi-Liquid Effects in Cyclotron Resonance . . . . . P. M. Platzman and K. C. Jacobs A974
Low-Field de Haas–van Alphen Effect in Copper . . . . . A. S. Joseph and A. C. Thorsen A979
Lattice Dynamics and Phase Transitions of Strontium Titanate . . . . . R. A. Cowley A981
Optical Absorption in an Electric Field . . . . . Joseph Callaway A998
Electron Spin Resonance on Interacting Donors in Silicon . . . . . D. Jérome and J. M. Winter A1001
Peierls Stress and Creep of a Linear Chain . . . . . J. H. Weiner and W. T. Sanders A1007
Incoherent and Thermal Small-Angle X-Ray Scattering in Silicon . . . . . D. L. Weinberg A1016
Optically Active Phonon Processes in CdS and ZnS . . . . . R. Marshall and S. S. Mitra A1019
Theory of Electron Backscattering . . . . . Roger F. Dashen A1025
Growing Helical Density Waves in Semiconductor Plasmas . . . C. E. Hurwitz and A. L. McWhorter A1033
Relationship between Neutron-Produced Prismatic Dislocations and the Change in Volume during Annealing in Copper . . . . . I. G. Greenfield and J. R. Demos A1051
Thermal Conductivity of Silicon and Germanium from 3°K to the Melting Point . . . . . C. J. Glassbrenner and Glen A. Slack A1058
Soft X-Ray Production by 1.5-MeV Protons . W. T. Ogier, G. J. Lucas, J. S. Murray, and T. E. Holzer A1070
Self-Compensation Limited Conductivity in Binary Semiconductors. I. Theory . . . . . G. Mandel A1073
Theory of the Optical and Magnetic Properties of the Self-Trapped Hole in Lithium Fluoride . . . . . T. P. Das, A. N. Jette, and R. S. Knox A1079
Electron Tunneling Through Thin Aluminum Oxide Films . . Thomas E. Hartman and Jay S. Chivian A1094
Interaction of Electron on the Donor Center with the Crystal Lattice in Germanium and Silicon . . . . . Andrzej Myszkowski and Stanisław Gómulka A1102
Measurement of the Conductivity Effective Mass in Semiconductors Using Infrared Reflection . . . . . Henry A. Lyden A1106
Pressure Effect on Resistivity of $Ga(As_{1-x}P_x)$ . . . . . G. E. Fenner A1113

LANCASTER PRESS, INC., LANCASTER, PA.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CONTENTS

THE PHYSICAL REVIEW SECOND SERIES, VOLUME 134, NO. 5A 1 JUNE 1964

Resistive Transition and Current Density Characteristics in Superconducting Niobium Containing Dissolved Gases . . . Warren DeSorbo A1119
Scattering of Energy-Time Wave Packets from Many-Body Systems . L. R. Dodd and I. E. McCarthy A1136
Coherence Properties of Blackbody Radiation. I. Correlation Tensors of the Classical Field . . . C. L. Mehta and E. Wolf A1143
Coherence Properties of Blackbody Radiation. II. Correlation Tensors of the Quantized Field . . . C. L. Mehta and E. Wolf A1149
Effect of Magnetic Field on Thermal Conductivity and Energy Gap of Superconducting Films . . . D. E. Morris and M. Tinkham A1154
Higher Order Coherent Raman Effects . . . C. L. Tang A1166
Interpretation of Frequency Shifts Due to Electron Exchange Collisions . . . P. L. Bender A1174
Fine Structure of Helium . . . Charles Schwartz A1181
Low-Energy Expansion of the Scattering Amplitude for Long-Range Quadrupole Potentials . . . Thomas F. O'Malley A1188
Lifetime of the First $^1P_1$ State of Zinc, Calcium, and Strontium . . . Allen Lurio, R. L. deZafra, and Robert J. Goshen A1198
g_J of the $(5s5p)$ $^3P_1$ Level of Cd and the $(6s6p)$ $^3P_1$ Level of Hg by High-Field Double Resonance . . . R. Kohler and P. Thaddeus A1204
Determination of the Polarizability Tensors of the Magnetic Substates of $^3P_2$ Metastable Argon . . . Edward Pollack, Edward J. Robinson, and Benjamin Bederson A1210
Studies of the Mechanism of Electron-Ion Recombination. II. . William A. Rogers and Manfred A. Biondi A1215
Ionization and Dissociation of Fast H_2 Molecules Incident on H_2 Gas . . . G. W. McClure A1226
Bound States in a Debye-Hückel Potential . . . Charles Ray Smith A1235
Slow Neutron Inelastic Scattering from Liquid Sodium . . . P. D. Randolph A1238
Theory of Solid Neon at 0°K . . . William J. Mullin A1249
Positronium in a LiH Crystal . . . S. M. Neamtan and R. I. Verrall A1254
Oscillatory Magnetoresistance in Gray Tin . . . E. D. Hinkley and A. W. Ewald A1261
$FeAl_2O_4$–$MgAl_2O_4$: Growth and Some Thermal, Optical, and Magnetic Properties of Mixed Single Crystals . . . Glen A. Slack A1268
Low-Temperature Galvanomagnetic Phenomena in an Intense Electric Field . . . Herbert F. Budd A1281
Spin-Lattice Relaxation Time of Conduction Electrons in Sodium Metal . . . Frederick Vescial, N. S. VanderVen, and R. T. Schumacher A1286
Absence of Pronounced Quadrupole Effects in the Nuclear Resonance of In^{115} in a Noncubic Environment . . . Lawrence H. Bennett and Rex J. Snodgrass A1290
Nuclear Magnetic Resonance in Lead Alloys . . . Rex J. Snodgrass and Lawrence H. Bennett A1294
Magnetic Properties of Nearly Free Electrons. Nonoscillatory Magnetic Susceptibility . . . M. L. Glasser A1296
Electron Paramagnetic Resonance or Fe^{3+} in the Strong Axial Field of $PbTiO_3$ Host . Denis J. A. Gainon A1300
Stimulated Phonon Emission by Supersonic Electrons and Collective Phonon Propagations . . . E. W. Prohofsky A1302
Optical Harmonic Generation in Single Crystal $BaTiO_3$ . . . Robert C. Miller A1313
Specific Heat of Niobium between 0.4 and 4.2°K . . . B. J. C. van der Hoeven, Jr. and P. H. Keesom A1320
Superconducting and Normal Specific Heats of a Single Crystal of Niobium . . . H. A. Leupold and H. A. Boorse A1322
Interstitials and Vacancies in α Iron . . . R. A. Johnson A1329
Electronic Spectra of Crystalline Germanium and Silicon . . . David Brust A1337
Redetermination of the Temperature Dependence of X Rays Anomalously Transmitted through Germanium . . . Boris W. Batterman A1354
Kinetic Secondary Electron Ejection from Tungsten by Cesium Ions . . . Samuel H. Bosch and Guntis Kuskevics A1356
Defects in Irradiated Silicon: Electron Paramagnetic Resonance and Electron-Nuclear Double Resonance of the Si-E Center . . . G. D. Watkins and J. W. Corbett A1359
Experimental Investigations of Single Injection in Compensated Silicon at Low Temperatures . . . Bob L. Gregory and Angel G. Jordan A1378
Ultrasonic Wave Propagation in Doped n-Germanium and p-Silicon . . . Warren P. Mason and T. B. Bateman A1387
Distribution Functions for the Number of Distinct Sites Visited in a Random Walk on Cubic Lattices: Relation to Defect Annealing . . . J. R. Beeler, Jr. A1396
Microscopic Dielectric Susceptibility of Germanium . . . W. Cochran and J. C. Phillips A1402

LANCASTER PRESS, INC., LANCASTER, PA.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CONTENTS

THE PHYSICAL REVIEW SECOND SERIES, VOLUME 134, NO. 6A 15 JUNE 1964

Bulk Superconductivity in Dilute Hexagonal Titanium Alloys . . . . . F. Heiniger and J. Muller A1407
Time-Reversed Motion in Kinetic Theory . . . . . M. B. Lewis A1410
Possibility of Synthesizing an Organic Superconductor . . . . . W. A. Little A1416
Cavitation in Liquid Helium . . . . . R. D. Finch, R. Kagiwada, M. Barmatz, and I. Rudnick A1425
Theory of an Optical Maser . . . . . Willis E. Lamb, Jr. A1429
Correlation Effects in Many Fermion Systems. II. Linked Clusters . . . . . Hugh P. Kelly A1450
Dynamics of a System of N Atoms Interacting with a Radiation Field . E. R. Buley and F. W. Cummings A1454
Collisional Electron Detachment from 20-MeV D^0 and D^- Ions . . . . . Klaus H. Berkner, Selig N. Kaplan, and Robert V. Pyle A1461
Hysteresis in Stability Conditions of Electron-Hole Plasma . . . . . Betsy Ancker-Johnson A1465
Stability of Longitudinal Oscillations in a Uniform Magnetized Plasma with Anisotropic Velocity Distribution . . . . . Lawrence S. Hall and Warren Heckrotte A1474
Lattice Vibrational Spectra of Beryllium, Magnesium, and Zinc . James A. Young and Juan U. Koppel A1476
Electron Spin Resonance Studies with Superconducting Magnets: The Spectrum of Dy^{3+} and Sm^{3+} in CaF_2 . . . . . W. Low A1479
Paramagnetic Resonance of Yb^{3+} in $CaWO_4$ . . . . . U. Ranon and V. Volterra A1483
Frequency Spectra of Body-Centered Cubic Lattices . . . . B. C. Clark, D. C. Gazis, and R. F. Wallis A1486
Paramagnetic Resonance and Relaxation of Trivalent Rare-Earth Ions in Calcium Fluoride. I. Resonance Spectra and Crystal Fields . . . . . M. J. Weber and R. W. Bierig A1492
Paramagnetic Resonance and Relaxation of Trivalent Rare-Earth Ions in Calcium Fluoride. II. Spin-Lattice Relaxation . . . . . R. W. Bierig, M. J. Weber, and S. I. Warshaw A1504
Ultrasonic Absorption at Microwave Frequencies and at Low Temperatures in MgO and Al_2O_3 . . . . . I. S. Ciccarello and K. Dransfeld A1517
Role of s-d Exchange Interactions in Dilute Alloys Exhibiting Both Low-Temperature Resistance Anomalies and Superconductivity . . . . . E. W. Collings, F. T. Hedgcock, and Y. Muto A1521
Covalency in Crystal Field Theory: $KNiF_3$ . . . . . R. E. Watson and A. J. Freeman A1526
Magnetic Moments and Unpaired Spin Densities in the Fe-Rh Alloys . . . . . G. Shirane, R. Nathans, and C. W. Chen A1547
General Theory of Cross Relaxation. I. Fundamental Considerations . . . . . W. J. C. Grant A1554
General Theory of Cross Relaxation. II. Higher Order Processes . . . . . W. J. C. Grant A1565
General Theory of Cross Relaxation. III. Application to Experiment . . . . . W. J. C. Grant A1574
Molecular Field Model and the Magnetization of YIG . . . . . Elmer E. Anderson A1581
Many-Body Theory of Localized d States in Metals. II. Pair of Impurities . . P. Gottlieb and H. Suhl A1586
Low-Temperature Magnetoresistance in Magnesium and Aluminum Containing Small Concentrations of Manganese or Iron . . . . . F. T. Hedgcock and Y. Muto A1593
F-Band Shape in the CsCl Structure Under Pressure . . . . . D. B. Fitchen A1599
Magnetic Translation Group . . . . . J. Zak A1602
Magnetic Translation Group. II. Irreducible Representations . . . . . J. Zak A1607
Nuclear Magnetic Relaxation of the Impurity Nucleus in Dilute Ferromagnetic Alloys . . R. L. Streever A1612
Further Results on the Fermi Surface of Beryllium . . . . . T. L. Loucks A1618
Specific Heat of Thulium Metal between 0.38 and 3.9°K . . . . . O. V. Lounasmaa A1620
Collective Strain Splitting of Acceptor States in Silicon . . . . . Calvin S. Koonce A1625
Band Structure of Gallium Phosphide from Optical Experiments at High Pressure . . . . . Richard Zallen and William Paul A1628
Thermal Disorder in a Dislocation Core and its Effect on Damping . . . . . P. D. Southgate and K. S. Mendelson A1642
Surface Effect in Secondary and Photoelectric Emission . . . . . I. Adawi A1649
Scattering of Spin Waves by Magnetic Defects . . . . . Joseph Callaway and Robert Boyd A1655

LANCASTER PRESS, INC., LANCASTER, PA.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

111678
Digitized by Arya Samaj Foundation Chennai and eGangotri
7032
111678

THE PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 134, No. 4A 18 MAY 1964

Resonant Frequencies of an Electromagnetic Cavity in an Accelerated System of Reference*

C. V. HEER
Department of Physics, Ohio State University, Columbus, Ohio
(Received 27 December 1963)

The resonant frequencies of an electromagnetic cavity in a nonpermanent gravitational field such as that provided by rotation is considered. The constitutive equations for accelerated macroscopic matter are developed and it is shown that the electromagnetic-field tensors are related by $H_{xt}=K_eF_{xt}$ and, $K_mH^{xy}=F^{xy}$, etc. For the special case of angular rotation about the z axis the constitutive equations are given in terms of noncovariant field vectors **B**, **H**, **D**, **E**. Using these constitutive equations Maxwell's equations are solved, and it is shown that the degeneracy between the axially degenerate modes, i.e., clockwise and counterclockwise traveling waves, is removed by the rotation. In vacuum due to the similarity between energy flux and momentum a shift in frequency of $\boldsymbol{\Omega}\cdot\mathbf{J}/\hbar$ or a splitting between traveling wave modes of $2\boldsymbol{\Omega}\cdot\mathbf{J}/\hbar$ is predicted. This led to a suggestion of the "Coriolis-Zeeman" effect for photons. Using either the above development or the energy density it is shown that the effect depends on the moment of the energy flux and is only proportional to the angular momentum. In the geometrical-optics region an index of refraction valid for accelerated macroscopic matter is developed and is applied to the square Fabry-Perot cavity rate gyroscope to yield the same result as either of the above methods. An experiment to show the validity of Minkowski's decomposition of the energy-momentum tensor is discussed.

INTRODUCTION

THE effect of acceleration on the resonant frequencies of an electromagnetic resonant structure or cavity is considered in this paper. Interference of light rays in a rotating system was suggested by Michelson as early as 1904 on the basis of the ether.[1] The observations of subsequent workers[2] confirmed the measurement of rotation by the interference technique and Michelson and Gale succeeded in measuring the angular rotation of the earth.[3] With the advent of the general theory of relativity it was shown that problems related to light transmission in accelerated frames of reference were more properly treated within the framework of general relativity. Adopting this viewpoint the author suggested that in rotating resonant-electromagnetic structures the degeneracy of the traveling wave modes about the axis of rotation would be removed by rotation and a beat frequency between the two modes of $2m\Omega$ would be observed.[4] This work was extended to optical masers[5] and confirmed by the experiments of Macek and Davis,[6] and Cheo and Heer.[7] It is the purpose of this paper to extend this earlier work.

In order to pursue the subject in a systematic manner the general principle of relativity that all systems of reference are equivalent with respect to the fundamental laws of physics is invoked. This requirement that the laws of nature be expressible in the form of equations which are form invariant or covariant is made possible by incorporating the metric $g_{\alpha\beta}$ into the physical laws. The $g_{\alpha\beta}$ are defined by the quadratic form

$$ds^2=g_{\alpha\beta}dx^\alpha dx^\beta. \quad (1)$$

For a system which is rotating with angular velocity Ω about the z axis and described by coordinates $x^1=x$, $x^2=y$, $x^3=z$, $x^4=ct$, the symmetric tensor $g_{\alpha\beta}$ has

* This work was supported by a grant from the National Science Foundation, and supported in part by the National Aeronautics and Space Administration through Grant NsG-552.

[1] A. A. Michelson, Phil. Mag. 8, 716 (1904).

[2] G. Sagnac, Compt. Rend. 157, 708 and 1410 (1913); F. Harress, dissertation, Jena, 1911 (unpublished); and B. Pogany, Ann. Physik 85, 244 (1928).

[3] A. A. Michelson and H. G. Gale, J. Astrophys. 61, 1401 (1925).

[4] C. V. Heer, Bull. Am. Phys. Soc. 6, 58 (1961).

[5] C. V. Heer, *Proceedings of the Third International Conference on Quantum Electronics* (Dunod Cie, Paris, and Columbia University Press, New York, 1963); A. H. Rosenthal, J. Opt. Soc. Am. 52, 1143 (1962).

[6] W. Macek and D. Davis, Appl. Phys. Letters 2, 67 (1963).

[7] P. K. Cheo and C. V. Heer (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar
Copyright © 1964 by The American Physical Society

Digitized by Arya Samaj Foundation Chennai and eGangotri

the following nonzero covariant and contravariant components.

$$\begin{aligned} &g_{xx}=g_{yy}=g_{zz}=1; && g_{xt}=g^{xt}=(-\Omega y/c);\\ &g_{yt}=g^{yt}=(\Omega x/c); && g^{xx}=(1-\Omega^2y^2/c^2);\\ &g^{yy}=(1-\Omega^2x^2/c^2); && g^{zz}=1; \qquad (2)\\ &g^{xy}=(\Omega^2xy/c); && g_{tt}=-[1-\Omega^2(x^2+y^2)/c^2];\\ &g^{tt}=-1; && \det g\equiv g=-1. \end{aligned}$$

Of the simple systems, i.e., constant linear velocity, linear acceleration, and angular rotation, only angular rotation does not permit the introduction of an instantaneous rest frame for the entire system and the metric cannot be placed in diagonal form. This aspect of the metric is considered in some detail by Möller.[8] It is also for this reason that effects of the order of v/c are observed in the rotating system. Thus, the experiment of Michelson can be explained by regarding the trajectory of a light ray as given by $ds^2=0$.

Maxwell's Equations

Following the procedure suggested in the Introduction, Maxwell's equations are written in a covariant form in an instantaneous rest frame, and by the rules of tensor algebra, remain valid in an arbitrary frame. For macroscopic matter Maxwell's equations in covariant form are

$$\frac{\partial F_{\alpha\beta}}{\partial x^\gamma}+\frac{\partial F_{\beta\gamma}}{\partial x^\alpha}+\frac{\partial F_{\gamma\alpha}}{\partial x^\beta}=0,$$

and

$$(-g)^{-1/2}\frac{\partial}{\partial x^\beta}[(-g)^{1/2}H^{\alpha\beta}]=j^\alpha. \qquad (3)$$

In the subsequent discussion $g_{\alpha\beta}$ remains dimensionless, $F_{\alpha\beta}$ and $H^{\alpha\beta}$ have the dimensions of (energy)$^{1/2}$, and the m.k.s. system of units is used. Although $F_{\alpha\beta}$ and $H^{\alpha\beta}$ can be written in terms of a polar vector and an axial vector in a covariant manner, a noncovariant notation[9] is found more convenient and consistent with conventional techniques. Nonconvariant vectors are defined as

$$\begin{aligned} &\epsilon_0^{1/2}E_x=F_{xt}; && \mu_0^{-1/2}B_x=\tfrac{1}{2}\delta_{xyz}F_{yz}; && \mu_0^{1/2}J_x=j^x;\\ &\epsilon_0^{-1/2}D_x=H^{tx}; && \mu_0^{1/2}H_x=\tfrac{1}{2}\delta_{xyz}H^{yz}; && \epsilon_0^{-1/2}\rho=j^t; \end{aligned} \qquad (4)$$

where δ_{xyz} is the three-dimensional Levi-Cevita symbol. Maxwell's equations take on their conventional appearance in this notation,

$$\begin{aligned} &\mathrm{curl}\mathbf{E}+\partial\mathbf{B}/\partial t=0, && \mathrm{div}\mathbf{B}=0,\\ &\mathrm{curl}\mathbf{H}-\partial\mathbf{D}/\partial t=\mathbf{J}, && \mathrm{div}\mathbf{D}=\rho, \end{aligned} \qquad (5)$$

but the operations curl and div are defined only for Cartesian ordering of the differentiation.

[8] C. Möller, *The Theory of Relativity* (Oxford University Press, London, 1957).
[9] J. Plebanski, Phys. Rev. 118, 1396 (1960).

Instantaneous Rest Frame

Using the rules of tensor algebra for two systems of coordinates which are related by

$$\begin{aligned} &dX^\Phi=\bar{A}^\Phi{}_\alpha dx^\alpha, && dx^\alpha=A^\alpha{}_\Phi dX^\Phi,\\ &\bar{A}^\Phi{}_\alpha A^\alpha{}_\Psi=\delta^\Phi{}_\Psi, && A^\alpha{}_\Psi\bar{A}^\Psi{}_\beta=\delta^\alpha{}_\beta, \end{aligned} \qquad (6)$$

the covariant and contravariant components of a tensor transform as

$$T_{\alpha\beta}=\bar{A}^\Phi{}_\alpha\bar{A}^\Psi{}_\beta T_{\Phi\Psi}, \quad \text{and} \quad T^{\alpha\beta}=A^\alpha{}_\Phi A^\beta{}_\Psi T^{\Phi\Psi}. \qquad (7)$$

Following the procedure discussed by Möller[8] an instantaneous rest frame is regarded as a system of inertia for which $G_{XX}=G_{YY}=G_{ZZ}=1$ and $G_{TT}=-1$. If (x,y,z,ct) and (X,Y,Z,cT) are regarded as instantaneously at rest with respect to each other, then

$$A^x{}_T=\bar{A}^X{}_t=0. \qquad (8)$$

Since $g_{\alpha\beta}$ is a tensor, the components of the metric $g_{\alpha\beta}$ are related to those in the instantaneous rest frame by

$$g_{\alpha\beta}=\bar{A}^\Phi{}_\alpha\bar{A}^\Psi{}_\beta G_{\Phi\Psi}. \qquad (9)$$

For subsequent reference,

$$g_{tx}=-\bar{A}^T{}_t\bar{A}^T{}_x, \quad \text{and} \quad g_{tt}=-\bar{A}^T{}_t\bar{A}^T{}_t. \qquad (10)$$

Also in the remainder of this paper capital Greek or Latin indices refer to the instantaneous rest frame, small indices refer to the frame under consideration, and repeated Greek letters are summed over the four indices. Repeated Latin letters are summed over only the space indices.

CONSTITUTIVE EQUATIONS

For vacuum $H^{\alpha\beta}=F^{\alpha\beta}$, and the relationship between **B**, **H**, **E** and **D** follow immediately from the transformation $F^{\alpha\beta}=g^{\alpha\mu}g^{\beta\gamma}F_{\mu\gamma}$. Macroscopic matter requires further elaboration. A procedure similar to that used by Minkowski for introducing the constitutive equations in special relativity is used to find the constitutive equations in a nonpermanent gravitational field. For isotropic dielectric and paramagnetic substances the constitutive equations in an instantaneous rest frame have their usual form

$$D_X=K_e\epsilon_0E_X, \quad B_X=K_m\mu_0H_X; \quad J_X=\sigma E_X \qquad (11)$$

in noncovariant notation. K_e is the dielectric constant, K_m the magnetic constant, and σ the electrical conductivity. In the instantaneous rest frame there is no distinction between covariant and contravariant components and these constitutive equations may be written as

$$\begin{aligned} &H_{XT}=K_eF_{XT}; \quad F_{YZ}=K_mH_{YZ};\\ &j_X=(\mu_0/\epsilon_0)^{1/2}\sigma F_{XT}; \quad j^T=\epsilon_0^{-1/2}\rho_I. \end{aligned} \qquad (12)$$

In any other frame of reference the components follow from the rules given by Eqs. (7) and (8). Thus

$$F_{xt}=\bar{A}^X{}_x\bar{A}^T{}_tF_{XT} \quad (\text{sum } X \text{ only}), \qquad (13a)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and similar expressions follow for H_{xt}. A simple relationship between H_{xt} and F_{xt} follows, i.e.,

$$H_{xt}=\bar{A}^{X}{}_{x}\bar{A}^{T}{}_{t}H_{XT}=K_{e}\bar{A}^{X}{}_{x}\bar{A}^{T}{}_{t}F_{XT}=K_{e}F_{xt}. \quad (13b)$$

Following a similar procedure for F^{xy} and H^{xy}, the first two constitutive equations may be written in the simple form as

$$H_{xt}=K_{e}F_{xt} \quad \text{and} \quad K_{m}H^{xy}=F^{xy}. \quad (14a)$$

Noncovariant notation may be introduced by relating the covariant and contravariant components;

$$K_{m}H^{xy}=F^{xy}=g^{x\alpha}g^{y\beta}F_{\alpha\beta}, \quad \text{and} \quad K_{e}F_{xt}=H_{xt}=g_{x\alpha}g_{t\beta}H^{\alpha\beta}.$$

For the special example of rotation along the z axis, the introduction of the metric given by Eqs. (2) and the noncovariant field vectors yields the following constitutive equations. With $\boldsymbol{\Omega}=\Omega\mathbf{a}_z$ these expressions reduce to

$$\mathbf{B}+\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\left[\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\mathbf{B}\right]=K_{m}\mu_{0}\mathbf{H}+c^{-1}\left[\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\mathbf{E}\right],$$
$$\mathbf{D}+\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\left[\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\mathbf{D}\right]=K_{e}\epsilon_{0}\mathbf{E}-c^{-1}\left[\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\times\mathbf{H}\right]. \quad (14b)$$

For vacuum, $K_e=1=K_m$, these equations may be rearranged to read

$$\mathbf{B}=\mu_{0}\mathbf{H}+(\mu_{0}/\epsilon_{0})^{1/2}[(\boldsymbol{\Omega}x\mathbf{r}/c)\times\mathbf{D}],$$
$$\mathbf{D}=\epsilon_{0}\mathbf{E}-(\epsilon_{0}/\mu_{0})^{1/2}[(\boldsymbol{\Omega}x\mathbf{r}/c)\times\mathbf{B}]. \quad (14c)$$

Rotation has the effect of making even the vacuum appear anisotropic. This effect occurs in first order of Ω and corresponds to the type term introduced by Michelson for the description in terms of the "aether."

The source equation transforms as a four vector, and by the rules of tensor algebra

$$j_{x}=\bar{A}^{X}{}_{x}j_{X}+\bar{A}^{T}{}_{x}j_{T}, \quad \text{and} \quad j_{t}=\bar{A}^{T}{}_{t}j_{T}. \quad (15)$$

For macroscopic matter the constitutive equations in the instantaneous rest system are given by Eqs. (12). Substituting for j_X and j_t yields

$$j_{x}=(-g_{tt})^{-1/2}(\mu_{0}/\epsilon_{0})^{1/2}\sigma F_{xt}-g_{tx}(-g_{tt})^{-1}j_{t}. \quad (16a)$$

Again using the metric to determine the components of j^x, the current and charge in the rotating system may be written as

$$\mathbf{J}=\sigma\left\{\mathbf{E}-\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\left[\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\cdot\mathbf{E}\right]\right\}\times[1-\Omega^{2}(x^{2}+y^{2})/c^{2}]^{-1/2},$$
$$\rho=\left\{\sigma c^{-1}\left(\frac{\boldsymbol{\Omega}x\mathbf{r}}{c}\right)\cdot\mathbf{E}+\rho_{I}\right\}[1-\Omega^{2}(x^{2}+y^{2})/c^{2}]^{-1/2}. \quad (16b)$$

Matter with a finite conductivity in the presence of an oscillating electric field will appear to have an induced oscillating current and charge. It is often desireable to separate the property of matter from the vacuum field. In the notation used in the paper this is expressed by

$$H^{\alpha\beta}=F^{\alpha\beta}-M^{\alpha\beta}$$

and noncovariant vectors $\mathbf{M}$ and $\mathbf{P}$ for the magnetization and polarization are defined as

$$\mu_{0}^{1/2}M_{x}=\tfrac{1}{2}\delta_{xyz}M^{xy}, \quad \text{and} \quad \epsilon_{0}^{-1/2}P_{x}=M^{xt}. \quad (17)$$

With these definitions, Maxwell's equations may be written in terms of $\mathbf{B}$ and $\mathbf{E}$ if $\mathbf{J}$ is replaced by

$$\mathbf{J}\rightarrow\mathbf{J}+\text{curl}\mathbf{M}+\partial\mathbf{P}/\partial t. \quad (18)$$

RESONANT FREQUENCY OF A ROTATING CAVITY

The approximate frequency of oscillation of a resonant electromagnetic cavity or structure is found by a method used by Slater[10] for microwave cavities. His method is modified to include traveling waves and it is necessary to use complex wave functions. A set of orthonormal functions $\mathbf{E}_a$ and $\mathbf{H}_a$ which are solutions of the vector Helmholtz equation and which satisfy the boundary conditions $\mathbf{n}\times\mathbf{E}_a=0$ and $\mathbf{n}\cdot\mathbf{H}_a=0$ on a perfectly conducting surface S and $\mathbf{n}\times\mathbf{H}_a=0$ and $\mathbf{n}\cdot\mathbf{E}_a=0$ on an insulating surface S' are introduced. These vectors are related by $-ik_a\mathbf{H}_a=\text{curl}\mathbf{E}_a$ and $ik_a\mathbf{E}_a=\text{curl}\mathbf{H}_a$. $k_a\mathbf{F}_a=\text{grad}\psi_a$ for the scalar potential. Orthogonality is chosen such that $\int dv\mathbf{E}_a\cdot\mathbf{E}_b^*=\delta(a=b)$, etc., and an arbitrary vector may be expanded in terms of this basic set, i.e.,

$$\mathbf{E}(\mathbf{r},t(=\sum_{a}\mathbf{E}_{a}\int dv\mathbf{E}\cdot\mathbf{E}_{a}^{*}+\sum_{\beta}\mathbf{F}_{\beta}\int dv\mathbf{E}\cdot\mathbf{F}_{\beta}^{*}, \quad (19)$$

where the volume integrals are the time-dependent coefficients. Expanding Maxwell's equations in this manner yields

$$-ik_{a}\int dv\mathbf{E}\cdot\mathbf{E}_{a}^{*}+\frac{d}{dt}\int dv\mathbf{B}\cdot\mathbf{H}_{a}^{*}=-\int_{S}dA(\mathbf{n}\times\mathbf{E})\cdot\mathbf{H}_{a}^{*}, \quad (20a)$$

$$+ik_{a}\int dv\mathbf{H}\cdot\mathbf{H}_{a}^{*}-\frac{d}{dt}\int dv\mathbf{D}\cdot\mathbf{E}_{a}^{*}=\int dv\mathbf{J}\cdot\mathbf{E}_{a}^{*}-\int_{S'}dA(\mathbf{n}\times\mathbf{H})\cdot\mathbf{E}_{a}^{*}, \quad (20b)$$

$$\int dv\mathbf{D}\cdot\text{grad}\psi_{a}^{*}+\int dv\rho\psi_{a}^{*}=0. \quad (20c)$$

This expansion in normal modes implies that the element

[10] J. C. Slater, *Microwave Electronics* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

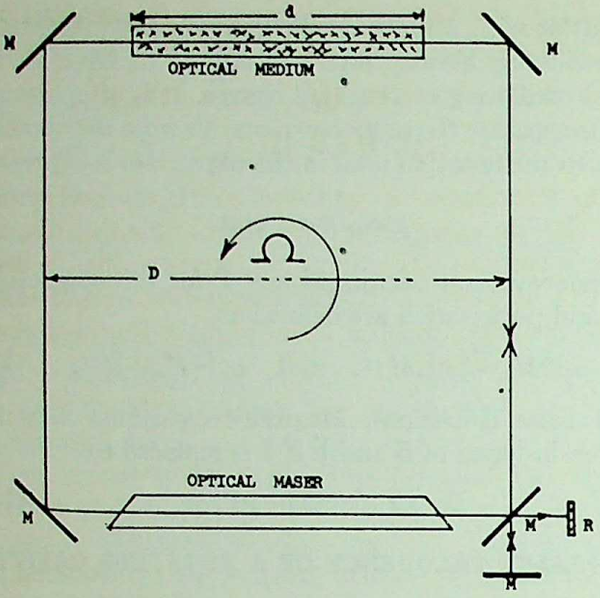


FIG. 1. Four mirrors M form the square Fabry-Perot cavity with diameter D. Optical media of length d and index of refraction $(K_eK_m)^{1/2}$ shifts the splitting in resonant frequency of the axially degenerate modes which is caused by the angular rotation Ω from the vacuum value given by Eq. (25) to that given by Eq. (26). Beats are observed with the photodetector R.

of volume is $dxdydz$ and the surface elements $dxdy$, etc. This is not true in the rotating system and a correction of the $\mathcal{O}(\Omega^2)$ is implied in the wave functions. The above equations are correct to $\mathcal{O}(\Omega^2)$ and should yield the resonant frequency to $\mathcal{O}(\Omega^2)$. Omitting the source terms on the right and using the constitutive equations to eliminate $\mathbf{B}$ and $\mathbf{D}$, the remaining equations may be solved to yield the resonant frequency. If $\nu_a{}^0$ is the resonant frequency, and $\mathbf{E}_a$ and $\mathbf{H}_a$ the normal modes for $\Omega=0$, the resonant frequency correct to $\mathcal{O}(\Omega)$ becomes

$$\nu_a=\nu_a{}^0\Big\{1+\tfrac{1}{2}(K_eK_mc^2)^{-1/2}\Omega \cdot\int dv[\mathbf{r}x(\mathbf{E}_a{}^*\times\mathbf{H}_a+\mathbf{E}_a\times\mathbf{H}_a{}^*)]\Big\}, \quad (21)$$

where as emphasized earlier $\boldsymbol{\Omega}$ is along the z axis. If two modes a and a' are axially degenerate for $\Omega=0$, then the frequency splitting between the two modes is given by

$$\Delta\nu/\nu_a=(K_eK_mc^2)^{-1/2}\Omega \cdot\int dv[\mathbf{r}x(\mathbf{E}_a\times\mathbf{H}_a{}^*+\mathbf{E}_a{}^*\times\mathbf{H}_a)]+\mathcal{O}(\Omega^2). \quad (22)$$

The cylindrical microwave cavity with rotation Ω along the z or symmetry axis provides an excellent example. In the TE_{mnp} mode the zero-order wave functions are of the form

$$\mathbf{E}_a=A_a\{\hat{a}_z\times\text{grad}[e^{im\varphi}Z_m(\alpha_{mn}\rho)]\}\sin\pi pz/z_0. \quad (23)$$

A direct evaluation of the integral for vacuum yields

$$\Delta\omega=2\pi\Delta\nu=2m\Omega, \quad (24)$$

and the rotation removes the twofold degeneracy of the axial mode.

A system of interest as a sensitive rate gyroscope[5–7] is the square Fabry-Perot cavity with confocal or flat mirrors. If $\mathbf{E}_a$ is an almost plane wave with a finite cross section, and if the perpendicular axis of the cavity makes angle γ with the z axis of rotation, integration over the beam volume yields a frequency splitting of

$$\Delta\nu/\nu_a=(\Omega D/c)\cos\gamma \quad (25)$$

for vacuum, D is the diameter of the square. Other configurations may be treated in a similar manner. If for convenience $\gamma=0$ and one arm of the Fabry-Perot contains matter as shown in Fig. 1 with an index of refraction $(K_eK_m)^{1/2}$ and of length d, then

$$\Delta\nu/\nu_a=(\Omega D/c)\{1-(d/4D)[1-(K_eK_m)^{-1/2}]\} \quad (26)$$

is the required correction for the matter.

Plane Waves

In the limit of short wavelengths or of the validity of geometrical optics the invariance of the phase $d\varphi=k_\alpha dx^\alpha$ and the invariance of the scalar $k_\alpha k^\alpha$ can be used to determine the index of refraction as measured by an observer in a nonpermanent gravitational field. Fermat's principle can be applied to determine the path. Using the transformation between k_t and k_T, the scalar invariant becomes

$$k_\Phi k^\Phi=k_\alpha k^\alpha=(\omega/c)^2(K_eK_m-1)(-g_{tt})^{-1}, \quad (27)$$

and is of course the usual null vector for vacuum. If $d\sigma$ is a line element along the path of the ray and n^a is the contravariant component of the ray direction, then

$$n^a=dx^a/d\sigma=k^ac\omega^{-1}(-g_{tt}/K_eK_m)^{1/2} \quad (28)$$

follows from Eq. (27) and the various relationships for the $g_{\alpha\beta}$.[11] Further manipulation yields for the phase

$$d\varphi=k_\alpha dx^\alpha=\omega c^{-1}\{[(-K_eK_m/g_{tt})^{1/2}-g_{at}n^a/g_{tt}]d\sigma-cdt\}. \quad (29)$$

If the quantity inside the square brackets is regarded as the index of refraction μ,

$$\mu=(-K_eK_m/g_{tt})^{1/2}-g_{at}n^a/g_{tt}, \quad (30)$$

then the path follows the variation

$$\delta\int\mu d\sigma=0. \quad (31)$$

Equation (31) is quite general and is Fermat's principle applicable for macroscopic matter and nonpermanent gravitational fields. If the principle of equivalence is invoked and for the $g_{at}=0$, the bending of a light ray in a macroscopic medium and in a permanent gravita-

[11] See L. Landau and E. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1951), p. 280, for a similar development for vacuum.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tional field described by g_{tt} is readily obtained. ($-g_{tt}=1+2\chi/c^2$ where χ is the gravitational potential.) It is interesting to note that the bending by a gravitational field is not separable from that bending by matter. For the special example a system rotating about the z axis and described by the metric given by Eq. (2), μ becomes

$$\mu=(K_eK_m)^{1/2}+c^{-1}(\boldsymbol{\Omega}\times\mathbf{r})\cdot\hat{n}+\mathcal{O}(\Omega^2), \tag{32}$$

where $\hat{n}$ is a unit normal in the direction of the ray. Macroscopic matter does not effect the term depending on g_{at} and the term of $\mathcal{O}(\Omega)$ in the index of refraction depends on the vacuum properties only. If this index of refraction is used for the square Fabry-Perot cavity, or cavities in which geometrical optics may be used with some validity, the shift in resonant frequency is the same as indicated by Eqs. (25) and (26).[12]

CORIOLIS-ZEEMAN EFFECT FOR PHOTONS

The removal of the twofold degeneracy of the axial mode of the cylindrical cavity discussed earlier yields an angular frequency of

$$\omega_{mnp}=\omega_{mnp}{}^0\pm m\Omega \quad (m=0,\ 1,\ 2,\ \cdots). \tag{33}$$

If the photon concept is used and $\mathbf{J}$ is the orbital angular momentum of the photon, the above equation may be written as

$$\hbar\omega=\hbar\omega^0+J_z\Omega, \tag{34}$$

and the correction is proportional to $\boldsymbol{\Omega}\cdot\mathbf{J}$. For vacuum, Eq. (22) is consistent with the concept of angular momentum. In fact, Eq. (25) is readily obtained by noting that

$$\Delta\nu/\nu=2\boldsymbol{\Omega}\cdot\mathbf{J}/h\nu=(\Omega D/c)\cos\gamma, \tag{35}$$

where $J_z=(h\nu/c)(D/2)\cos\gamma$ is the projection of the orbital angular momentum $\mathbf{J}$ along the z axis. The observer in the rotating system would find the resonant frequency dependent on the angular momentum of the photon and would regard the level splitting as a "Coriolis-Zeeman" effect for the photon. This concept is valid only for vacuum. The right side of Eq. (22) is the expression for the moment of energy flux rather than angular momentum. This distinction between energy flux and momentum is apparent only in the presence of matter or in $\mathcal{O}(\Omega^2)$.

If the energy-momentum tensor is defined as

$$S_\alpha{}^\beta=F_{\alpha\mu}H^{\beta\mu}-\tfrac{1}{4}\delta_\alpha{}^\beta F_{\mu\nu}H^{\mu\nu}, \tag{36}$$

whose covariant divergence yields the four-force $f_\alpha=F_{\alpha\beta}j^\beta$, the various components in the noncovariant notation are

$$\begin{aligned} &S_t{}^x=-c^{-1}(\mathbf{E}\times\mathbf{H})_x, \quad S_x{}^t=c(\mathbf{D}\times\mathbf{B})_x, \\ &S_t{}^t=-\tfrac{1}{2}(\mathbf{E}\cdot\mathbf{D}+\mathbf{B}\cdot\mathbf{H}). \end{aligned} \tag{37}$$

In the instantaneous rest frame $S_T{}^T$ is the energy density, $S_T{}^X$ the energy flux, and $S_X{}^T$ the momentum, and these quantities have the same interpretation in other systems of coordinates. The energy-momentum tensor in the instantaneous rest frame and the frame of observation are related by

$$S_\Phi{}^\Psi=A^\alpha{}_\Phi\bar{A}^\Psi{}_\beta S_\alpha{}^\beta, \tag{38}$$

and the energy density may be written as

$$S_T{}^T=S_t{}^t+(g_{tx}/g_{tt})S_t{}^x. \tag{39}$$

For the stationary system which is being considered the $g_{\alpha\beta}$ do not depend on the local time variable t and for short periods of time in which losses may be ignored, solutions of Maxwell's equations which are independent of the time variable may be obtained. The observer in the moving system interprets the integral of $S_t{}^t$ over the volume of the cavity or radiation field at time t as a constant of motion for the system. If a single photon is present the photon frequency is measured as

$$\nu=h^{-1}\int dv\, S_t{}^t=h^{-1}\int dv[S_T{}^T-(g_{tx}/g_{tt})S_t{}^x]. \tag{40}$$

Considered as a traveling wave the integral over $S_T{}^T$ is independent of the sense of rotation and the beat frequency caused by the interference of clockwise (c.w.) and counter-clockwise (c.c.w.) photons at a detector attached to the moving system is

$$\begin{aligned} \Delta\nu&=2h^{-1}\int dv[g_{tx}/g_{tt}]S_t{}^x \\ &=2h^{-1}c^{-2}\boldsymbol{\Omega}\cdot\int dv[\mathbf{r}x(\mathbf{E}\times\mathbf{H})]+\mathcal{O}(\Omega^2). \end{aligned} \tag{41}$$

This is the same as the earlier expression given by Eq. (22) when $\mathbf{E}$ and $\mathbf{H}$ are replaced by zero-order orthonormal $\mathbf{E}_a$ and $\mathbf{H}_a$ wave functions. The dependence on the projection of the moment of the energy flux rather than the angular momentum is immediately apparent.

MINKOWSKI'S ENERGY-MOMENTUM TENSOR

The lack of symmetry in the Minkowski energy-momentum tensor between the energy flux $S_T{}^\Phi$ and the momentum $S_\Phi{}^T$ in the presence of macroscopic matter has been discussed frequently in the literature.[13]

The energy-momentum tensor is symmetric in an inertia frame in the absence of matter and becomes

[12] Although an experiment with matter waves does not seem possible at this time, it is interesting to note that using the same formalism with the Klein-Gordon equation a phase difference between the c.w. and c.c.w. paths is obtained in which the Compton wavelength $\lambda=h/mc$ replaces the optical wavelength. The phase shift is $8\pi D^2\Omega/\lambda c$ in an experiment similar to that used by Michelson (Ref. 3). Interference depends on the Compton rather than the de Broglie wavelength (Ref. 4).

[13] See Ref. 8, p. 204 for a discussion and references to related discussions.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

nonsymmetric in the presence of matter or nonpermanent gravitational fields. A first-order nonsymmetry appears in the presence of rotation. Minkowski's theory requires that transparent matter not exchange energy with the electromagnetic field. It is shown in the earlier developments in this paper that the frequency shift in first order of Ω depends on the moment of the energy flux and also that the change in index of refraction in the geometrical optic region is independent of the presence of matter to $\mathcal{O}(\Omega)$. These same results follow with the Minkowski energy-momentum tensor given by Eq. (36). Thus an observation of the beat frequency in accord with Eq. (26) in the presence of matter would support the Minkowski formulation. This effect should be observable in the square Fabry-Perot cavity with flat mirrors for which the mode of operation is simple. Since the presence of the maser material allows the cavity to oscillate at both frequencies and the beat frequency is sensitive to rotation only, the shift in the beat frequency is measureable. Combining Eqs. (25) and (26) for ν_a coinciding with a maser line, the fractional shift is

$$(\Delta\nu_M/\Delta\nu_0)=1-(d/4D)[1-(K_eK_m)^{-1/2}], \quad (42)$$

where $\Delta\nu_M$ is the beat frequency with matter and $\Delta\nu_0$ with vacuum. If a large portion of the maser is filled with matter an effect as large as 30% can be expected. Unfortunately terms of $\mathcal{O}(\Omega^2)$ cannot be tested by present experimental apparatus. Applying the principle of equivalence it is the terms of $\mathcal{O}(\Omega^2)$ which are equivalent to the effects of a permanent gravitational field. Although not emphasized in the development, the equations developed in this paper may be applied in a permanent gravitational field by invoking the principle of equivalence.

EXCITATION OF CAVITY MODES

Using the form invariance of the fundamental laws of physics and the concept of the instantaneous rest frame, the constitutive equations for macroscopic matter in an accelerated system have been developed and used with Maxwell's equations to discuss the resonant frequencies of electromagnetic cavities. The deformation of matter by the acceleration forces has been neglected. Since the deformation of matter by the acceleration would change the shape of the cavity and the physical constants of enclosed matter, the small shifts suggested would be difficult to observe. For axially degenerate modes the deformations affect both c.w. and c.c.w. traveling waves in the same manner and the splitting between modes given by Eq. (22) is almost independent of the deformations. The splitting between modes given by Eq. (22) is valid over the entire region of the electromagnetic spectrum for which cavity modes may be defined. Thus, in the microwave region a reasonable size cavity would have a value of $m\sim10$, while in the optical region a value of $m>10^6$ is possible. For a closed microwave cavity the normal modes may be determined in a standard manner and wall losses introduced by discussing the decay of stored energy in terms of the Q. Since the beat frequency would be of the order of cycles per second for a rotation rate of 1 rad per sec, a Q of the order of 10^9 would be required to resolve the splitting between the c.w. and c.c.w. modes. Such a Q is almost possible with superconducting cavity walls. In the optical region the Fabry-Perot cavity shown in Fig. 1 may be discussed in a similar manner. A system with dimensions of the order of 1 m has the Q limited to 10^8 by diffraction[14] and reflection losses and a beat note of the order of kilocycles is not resolved.

In order to enhance the apparent Q of the system a maser material in the appropriate frequency range may be used.[5–7] The volume source or sink term on the right-hand side of Eq. (20b) may be due to either the polarization, magnetization, or conductivity as shown by Eq. (18). If the maser media has a narrow linewidth described by $\epsilon''(\nu)$ or $\chi''(\nu)$,[15] either property may be discussed in terms of a generalized negative resistance $\sigma(\nu)$ which is proportional to it. In this linear approximation the response of the c.w. mode depends on the value of $\sigma(\nu)$ at the frequency of the c.w. mode. A similar argument follows for the c.c.w. mode, and both modes will be excited for a maser linewidth greater than the splitting. Oscillation will occur for both modes in the linear approximation. If the polarization depends on the electric field in a nonlinear manner, the question of coupling between these closely spaced modes needs further consideration.[16] It may be possible to study these nonlinearities by using the square Fabry-Perot cavity as a precision spectrometer system.

ACKNOWLEDGMENTS

The author wishes to acknowledge many helpful discussions with Dr. R. L. Mills, Dr. J. R. Gaines, and P. K. Cheo. The name "Coriolis-Zeeman" effect for a photon was suggested by a participant at the New York meeting of the American Physical Society.[4]

[14] A. G. Fox and T. Li, Bell System Tech. J. **40**, 453 (1961).
[15] J. P. Gordon, H. Z. Zeiger, and C. H. Townes, Phys. Rev. **99**, 1264 (1955); A. Javan, W. R. Bennett, Jr. and D. R. Herriott, Phys. Rev. Letters **6**, 106 (1961); W. R. Bennett, Jr., Appl. Opt. Suppl. p. 24 (1962), for a discussion and a list of references.
[16] W. E. Lamb, Jr. (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Theory of the Electronic Thermal Conductivity of Superconductors with Strong Electron-Phonon Coupling

VINAY AMBEGAOKAR*
Department of Physics and Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York

AND

LUDWIG TEWORDT†
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received 7 October 1963; revised manuscript received 13 January 1964)

The electronic thermal conductivity of superconductors is discussed without recourse to the quasiparticle approximation. The Kubo formula for the thermal conductivity is used as the starting point. This is first examined in the Hartree-Fock approximation in the Nambu space. It is shown that in the Eliashberg approximation of neglecting the momentum dependence of the electronic self-energy, the calculation of the thermal conductivity is reduced to a quadrature, involving however the complex energy gap and renormalization functions which are solutions of the Eliashberg gap equations at finite temperatures. These equations are given in an Appendix. The problem is also considered in the ladder approximation in the Nambu space, and a generalized Boltzmann equation is derived which includes corrections to the Hartree-Fock approximation corresponding to the replacement of the scattering by the transport cross section. It is shown that the standard Boltzmann equation for superconducting quasiparticles is obtained in the weak-coupling limit. No numerical calculations are performed in this paper, but a clear scheme for such calculations is outlined. Reasons for believing that such calculations will explain the anomalous drop in the electronic thermal conductivity of superconducting lead are given.

I. INTRODUCTION

EXPERIMENTS on the strong elemental superconductors lead and mercury show that the electronic thermal conductivity of these materials differs markedly from that of typical weak superconductors like tin or indium. For lead and mercury[1] the ratio, K_s/K_n, of the thermal conductivity in the superconducting and normal states when plotted against the reduced temperature T/T_c shows a steep positive slope of about 5 near $T=T_c$. For tin[2] and indium,[3] on the other hand, the experiments yield a smaller slope of about 1.6.

A theory of the electronic thermal conductivity of superconductors based on the quasiparticle approximation and the Boltzmann equation approach of Bardeen, Rickayzen, and one of the present authors[4] has been carried through previously.[5] The results obtained are in substantial agreement with the data on tin and indium. The lack of agreement between the theory given in I and the data on lead is not surprising since, as shown in the work of Schrieffer, Scalapino, and Wilkins,[6] the strong electron-phonon coupling causes the quasiparticle picture to be quite meaningless over much of the energy spectrum.

It therefore seems reasonable to study the thermal conductivity of lead on the basis of the strong-coupling theory used by Schrieffer *et al.*[6] in explaining tunneling characteristics. This is the aim of the present paper. We use the method of thermodynamic Green's functions[7] and thereby avoid the quasiparticle approximation. We expect that in the limit of almost stable excitations our theory will reduce to the theory developed in I, and this expectation is realized. In this paper we carry the theory as far as we are able to without using numerical methods. We expect to use the formulas obtained here as the basis for further calculation.

Our starting point is the Kubo formula in which the thermal conductivity is expressed in terms of the correlation function of two heat current operators.[8] For the superconducting state it is convenient to use the two component space introduced by Nambu.[9] In Sec. II the problem is treated in the Hartree-Fock approximation which means that the correlation function is factored into two Nambu matrix Green's functions. The momentum integration in the expression for the thermal

---

* Alfred P. Sloan Foundation Fellow.

† Permanent address: Institut für Angewandte Physik, Universität Hamburg, Hamburg, Germany.

[1] J. K. Hulm, Proc. Roy. Soc. (London) A204, 98 (1956); J. H. P. Watson and G. M. Graham, Can. J. Phys. 41, 1738 (1963).

[2] A. M. Guenault, Proc. Roy. Soc. (London) A262, 420 (1961).

[3] A. M. Toxen and R. E. Jones, Phys. Rev. 120, 1167 (1960).

[4] J. Bardeen, G. Rickayzen, and L. Tewordt, Phys. Rev. 113, 982 (1959).

[5] L. Tewordt, Phys. Rev. 129, 657 (1963), hereafter referred to as I. Other work on this subject is contained in L. Tewordt, *ibid.* 128, 12 (1962), hereafter referred to as II; L. P. Kadanoff and P. Martin, *ibid.* 124, 678 (1961); B. T. Geilikman, V. Z. Kresin, Zh. Eksperim. i Teor. Fiz. 41, 1142 (1961) [English transl.: Soviet Phys.—JETP 14, 816 (1962)].

[6] J. R. Schrieffer, D. J. Scalapino, and J. W. Wilkins, Phys. Rev. Letters 10, 336 (1963).

[7] Several references on the Green's function method are now available. A reference in which the treatment is similar in style and notation to that of the present paper is V. Ambegaokar, *1962 Brandeis Lectures* (W. A. Benjamin, Inc., New York, 1963), Vol. 2. Other references are given here.

[8] We use the form given by J. S. Langer, Phys. Rev. 128, 110 (1962). See also Ref. 7.

[9] Y. Nambu, Phys. Rev. 117, 648 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

conductivity can be carried out if one makes the weak-momentum approximation introduced by Eliashberg,[10] that is, if one neglects the $\hat{p}$ dependence of the renormalization function $Z_p(\omega)$ and the gap function $\Delta_p(\omega)$. In contrast to the result of the quasiparticle approach, the thermal conductivity is expressed as an integral over all excitation energies rather than over all momenta. The integrand contains the complex functions $Z(\omega)$ and $\Delta(\omega)$. These functions have to be determined from the generalization to finite temperatures of Eliashberg's gap equations. The generalized gap equations are given in Appendix A. In the limit of zero temperature these equations have been solved for lead by Schrieffer *et al.* However, for the problem of the thermal conductivity one needs the solutions near T_c.

In Sec. III the problem is treated in the more general ladder approximation in the Nambu space. This is the approximation that must be used with the superconducting approximation to the single-particle self-energy operator in order to preserve conservation laws[11] or, equivalently, Ward identities. To facilitate the treatment of this approximation we use a method introduced recently by Langer, Maradudin, and Wallis[12] in discussing optical absorption by anharmonic crystals. The integral equation for the vertex part is simplified by using a double spectral representation (derived in Appendix B) for the vertex function. The resulting integral equation for one of the spectral functions corresponds to a generalized Boltzmann equation in which this function essentially plays the role of the non-equilibrium part of the distribution function. The integral equation simplifies if one makes an Ansatz for the vector part of the spectral function which corresponds to Bloch's Ansatz in the elementary theory, and further, if one again uses the weak-momentum approximation. In this way one obtains an integral equation in the energy variable alone which represents the generalization of Bloch's equation for thermal conductivity to a strongly coupled superconductor. In the limit of weak electron-phonon interaction the equation reduces to the Boltzmann equation used in I.

II. HARTREE-FOCK APPROXIMATION

If we neglect the entropy flow and measure energies with respect to the chemical potential,[13] the Kubo formula for the thermal conductivity K may be written as follows (we use units in which $\hbar=1$):

$$K=\frac{2}{3VT}\,\mathrm{Im}\int_{-\infty}^{0}dt_2t_2\int d^3x_1d^3x_2\langle\mathbf{u}(\mathbf{x}_1,0)\cdot\mathbf{u}(\mathbf{x}_2,t_2)\rangle. \quad (2.1)$$

Above V is the volume of the system, T is the temperature, and $\mathbf{u}$ is the heat-current density operator at the space-time point $1=(\mathbf{x}_1,t_1)$. The brackets denote an average in the grand canonical ensemble. For the interacting electron-phonon system $\mathbf{u}$ has the form $\mathbf{u}=\mathbf{u}_{\mathrm{el}}+\mathbf{u}_{\mathrm{ph}}$, where

$$\mathbf{u}_{\mathrm{el}}=-\frac{1}{2m}\sum_\sigma\left(\frac{\partial}{\partial t_1}\nabla_{1'}+\frac{\partial}{\partial t_{1'}}\nabla_1\right)\psi_\sigma^\dagger(1')\psi_\sigma(1)\,|_{1'=1} \quad (2.2)$$

is the heat current for the electronic system. It includes that part of the electron-phonon interaction energy which is due to electronic charge fluctuations. [The operators $\psi_\sigma^\dagger(1')$ and $\psi_\sigma(1)$ create and annihilate electrons with spin σ at $1'$ and 1, respectively.]

The detailed form of the phonon part of the heat-current density operator[14] is not needed for the purposes of this paper. It suffices to note that the correlation function of $\mathbf{u}_{\mathrm{ph}}$ with itself gives rise through Eq. (2.1) to a thermal conductivity which may be interpreted as the thermal conductivity of the lattice as limited by electron scattering.[15] In the normal state this contribution is proportional to T^2 and is too small to be observed in a pure system. In the superconducting state the lattice thermal conductivity is increased relative to that in the normal state because the energy gap in the electronic spectrum leads to an increase in the relaxation time for phonons. However, because of the smallness of the lattice term, the electronic contribution (though rapidly falling) will dominate the thermal conductivity just below T_c which is the region of interest to us. The cross terms in the expression (2.1) between $\mathbf{u}_{\mathrm{el}}$ and $\mathbf{u}_{\mathrm{ph}}$ describe "phonon drag" effects which are totally negligible for heat conduction because the electronic heat current is odd with respect to reflections in the Fermi surface. In what follows, therefore, we take (2.2) as the expression for the heat current operator and drop the subscript "el." In the language of conventional transport theory we are assuming that the lattice is always in equilibrium. The contribution to the thermal conductivity of the processes by which the lattice does in fact achieve equilibrium are, as we have seen, small in the region of interest to us.

Let us introduce the correlation function

$$P(1,2)=\langle T[\mathbf{u}(1)\cdot\mathbf{u}(2)]\rangle, \quad (2.3)$$

where T is Wick's time ordering operator. In the standard way we make Fourier transformations in the spatial and imaginary time difference coordinates, so

[10] G. M. Eliashberg, Zh. Eksperim. i Teor. Fiz. 38, 960 (1960) [English transl. Soviet Phys.—JETP 9, 1385 (1959)].

[11] G. Baym and L. P. Kadanoff, Phys. Rev. 124, 287 (1961).

[12] J. S. Langer, A. A. Maradudin, and R. F. Wallis, Proceedings of the International Conference on Lattice Dynamics, 1963 (to be published).

[13] See L. P. Kadanoff and P. C. Martin, Ann. Phys. (N. Y.) 24, 419 (1963), for a discussion of this point.

[14] In a continuum Debye approximation this has the form given by K. Baumann, Ann. Phys. (N. Y.) 23, 221 (1963).

[15] Because the phonon lifetimes contain the small quantity $(m/M)^{1/2}$ where M is the ionic mass, the quasiparticle approximation is valid for the phonons. {See Ref. 10 and A. B. Migdal, Zh. Eksperim. i Teor. Fiz. 34, 1438 (1958) [English transl.: Soviet Phys.—JETP 14, 816 (1962)].} By using the methods of the present paper one can recover a Boltzmann equation for the thermal conduction by phonons as limited by electron scattering.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that

$$P(1,2)=\int\frac{d^3q}{(2\pi)^3}\frac{i}{\beta}\sum_{\nu_m}P(q,\nu_m)\times\exp\{i[\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)-\nu_m(t_1-t_2)]\}. \quad (2.4)$$

Here $\nu_m=2\pi mi/\beta$, m runs over all integers, and $\beta=(k_BT)^{-1}$ with k_B Boltzmann's constant. In terms of $P(q,\nu_m)$ Eq. (2.1) becomes

$$K=-\frac{1}{3T}\frac{d}{d\nu}\{\mathrm{Re}[P(q,\nu_m)|_{\nu_m=\nu-i0^+}]\}|_{\substack{\nu=0\\q=0}}. \quad (2.5)$$

It is convenient to use Nambu's two-component field operators defined by

$$\psi(1)=\begin{pmatrix}\psi_\uparrow(1)\\ \psi_\downarrow^\dagger(1)\end{pmatrix},\quad \psi^\dagger(1)=(\psi_\uparrow^\dagger(1)\psi_\downarrow(1)).$$

In terms of these and the Pauli spin matrices τ_i the correlation function may be written as follows

$$P(1,2)=\frac{1}{4m^2}\left(\frac{\partial}{\partial t_1}\nabla_1'+\frac{\partial}{\partial t_1'}\nabla_1\right)\cdot\left(\frac{\partial}{\partial t_2}\nabla_2'+\frac{\partial}{\partial t_2'}\nabla_2\right)(\tau_3)^{ij}(\tau_3)^{kl}\times\langle T[\psi_i(1)\psi_k(2)\psi_l^\dagger(2')\psi_j^\dagger(1')]\rangle|_{\substack{1'=1^+\\2'=2^+}}. \quad (2.6)$$

In writing the correlation function in the form (2.6) we have neglected certain discontinuity terms that arise from the time derivatives acting on the time ordering operator. These terms make no contribution to the thermal conductivity.

In this section we shall derive the expression for K in the Hartree-Fock approximation. This is defined by

$$\langle T[\psi_i(1)\psi_k(2)\psi_l^\dagger(2')\psi_j^\dagger(1')]\rangle\rightarrow\mathcal{G}_{il}(12')\mathcal{G}_{kj}(21'), \quad (2.7)$$

where $\mathcal{G}_{il}$ is Nambu's matrix Green's function

$$\mathcal{G}_{il}(12)=-i\langle T[\psi_i(1)\psi_l^\dagger(2)]\rangle. \quad (2.8)$$

Then we obtain from Eqs. (2.6), (2.7), and (2.4)

$$P(q,\nu_m)=\frac{1}{4m^2}\int\frac{d^3p}{(2\pi)^3}\frac{i}{\beta}\sum_{\zeta_l}e^{\zeta_l0^+}[\zeta_l(\mathbf{p}+\mathbf{q})+(\zeta_l+\nu_m)\mathbf{p}]^2\times\mathrm{Tr}\{\tau_3\mathcal{G}(p+q,\zeta_l+\nu_m)\tau_3\mathcal{G}(p,\zeta_l)\}, \quad (2.9)$$

where $\zeta_l=(2l+1)\pi i/\beta$ and l takes on integral values. The Hartree-Fock approximation for the heat-current correlation function is represented by the diagram in Fig. 1.

We now insert the spectral representation for $\mathcal{G}(p,\zeta_l)$,

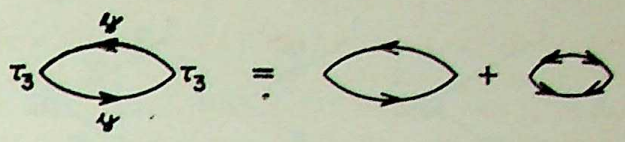


FIG. 1. Diagrammatic representation of the Hartree-Fock approximation for the heat-current correlation function of superconductors. The solid lines on the left of the equality sign represent Nambu matrix propagators $\mathcal{G}$. The equivalent diagrammatic representation in terms of Gorkov's normal and anomalous Green's functions is shown on the right. In the present formalism the terms on the right arise from taking the trace $\mathrm{Tr}[\tau_3\mathcal{G}\tau_3\mathcal{G}]$.

namely,

$$\mathcal{G}(p,\zeta_l)=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{a(p,\omega)}{\zeta_l-\omega} \quad (2.10)$$

into (2.9) and carry out the sum over ζ_l. Introducing the result into (2.5) one finds

$$K=\frac{1}{24m^2k_BT^2}\int\frac{d^3pd\omega}{(2\pi)^4}p^2\omega^2\,\mathrm{sech}^2(\tfrac{1}{2}\beta\omega)\times\mathrm{Tr}\{\tau_3a(p,\omega)\tau_3a(p,\omega)\}. \quad (2.11)$$

The Nambu matrix propagator $\mathcal{G}(p,\zeta_l)$ with ζ_l continued to the complete complex z plane has the form

$$\mathcal{G}(p,z)=\frac{zZ_p(z)+\epsilon_p\tau_3+\phi_p(z)\tau_1}{z^2Z_p{}^2(z)-\epsilon_p{}^2-\phi_p{}^2(z)}, \quad (2.12)$$

where $\epsilon_p=(p^2/2m)-\mu$. The renormalization function $Z_p(z)$ and the gap function $\phi_p(z)$, both complex, have to be determined from the Nambu-Eliashberg self-consistency condition. We now make the weak-momentum approximation introduced by Eliashberg, that is we neglect the p dependence of $Z_p(z)$ and $\phi_p(z)$. Then the integration over p, or ϵ_p, in Eq. (2.11) can be done. The spectral function $a(p,\omega)$ is obtained from (2.12) from the relation

$$a(p,\omega)=2\,\mathrm{Im}\mathcal{G}(p,\omega-i0^+)=(1/i)[\mathcal{G}(p,\omega-i0^+)-\mathcal{G}(p,\omega+i0^+)]. \quad (2.13)$$

The integrals over ϵ that are then needed are

$$I_1=\int_{-\infty}^{\infty}d\epsilon\,\mathrm{Tr}\{\tau_3\mathcal{G}(p,\omega-i0^+)\tau_3\mathcal{G}(p,\omega-i0^+)\}=0, \quad (2.14)$$

and

$$I_2=\int_{-\infty}^{\infty}d\epsilon\,\mathrm{Tr}\{\tau_3\mathcal{G}(p,\omega-i0^+)\tau_3\mathcal{G}(p,\omega+i0^+)\}=\frac{\pi}{|\mathrm{Im}\{Z(\omega)[\omega^2-\Delta^2(\omega)]^{1/2}\}|}\times\left(1+\frac{\omega^2-|\Delta(\omega)|^2}{|\omega^2-\Delta^2(\omega)|}\right). \quad (2.15)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

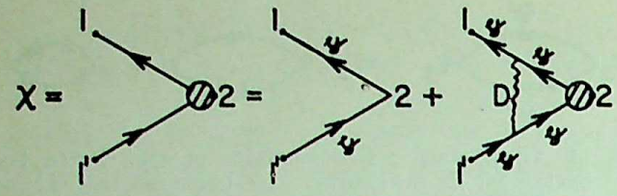


FIG. 2. Integral equation for the vertex function in the ladder approximation, represented in terms of diagrams. $\mathcal{G}$ and D denote particle and phonon propagators, respectively.

The integrations are straightforward and are sketeched in Appendix C. The complex functions $Z(\omega)$ and $\Delta(\omega)$, the latter being the complex energy gap parameter, are related to the functions occurring in (2.12) by

$$Z(\omega)=Z_1(\omega)+iZ_2(\omega)=Z(z=\omega-i0^+);$$
$$\Delta(\omega)=\Delta_1(\omega)+i\Delta_2(\omega) \qquad (2.16)$$
$$=\phi(z=\omega-i0^+)/Z(z=\omega-i0^+).$$

The thermal conductivity K is now given by

$$K=\frac{n}{8mk_BT^2}\int_0^\infty d\omega\frac{\omega^2\,\mathrm{sech}^2(\tfrac{1}{2}\beta\omega)}{|\mathrm{Im}\{Z(\omega)[\omega^2-\Delta^2(\omega)]^{1/2}\}|}\times\left(1+\frac{\omega^2-|\Delta(\omega)|^2}{|\omega^2-\Delta^2(\omega)|}\right). \qquad (2.17)$$

Above n is the density of electrons. Near $T=0$ the factor in the large brackets becomes zero for $\omega<\Delta_0$, where $\Delta_0=\Delta_1(\Delta_0)$ is the gap edge, since Δ_2 vanishes for $\omega<\Delta_0$. At finite temperatures the existence of thermally excited phonons may cause Δ_2 to have a small nonzero value for $\omega<\Delta_0$.

If for those frequencies ω that contribute predominantly to the integral in (2.17) the inequality $2\Delta_2\Delta_1<\omega^2-\Delta_1^2+\Delta_2^2$ is satisfied, the following approximate formula for K obtains

$$K=\frac{n}{8mk_BT^2}\int_0^\infty d\omega\frac{\omega^2(\omega^2-\Delta_1^2+\Delta_2^2)^{1/2}}{|Z_2(\omega^2-\Delta_1^2+\Delta_2^2)-Z_1\Delta_1\Delta_2|}\times\mathrm{sech}^2(\tfrac{1}{2}\beta\omega)\left(1+\frac{\omega^2-\Delta_1^2-\Delta_2^2}{|\omega^2-\Delta_1^2+\Delta_2^2|}\right). \qquad (2.18)$$

To first order in Δ_2 and Z_2 the expression in the denominator of the integrand in (2.18) reduces to $\omega Z_1\Gamma$, where Γ is one-half the decay rate for a quasiparticle as calculated in II. Equation (2.18) is thus somewhat more general than the quasiparticle approximation. In the case of lead it would appear that the condition $\Delta_2\Delta_1<\omega^2-\Delta_1^2+\Delta_2^2$ is reasonably well satisfied for frequencies below 2 or $3k_BT_c$, the latter being roughly the frequency which contributes maximally to K, but in a numerical calculation it is probably no harder to use the exact form (2.17).

III. GENERALIZED BOLTZMANN EQUATION

We now wish to obtain a Boltzmann-like equation for the correlation function $P(1,2)$. It is convenient first to rewrite the expression for this quantity as follows:

$$P(1,2)=\frac{1}{2m}\left(\frac{\partial}{\partial t_1}\nabla_{1'}+\frac{\partial}{\partial t_{1'}}\nabla\right)\cdot\mathrm{Tr}\{\tau_3\chi(1,2,1')\}|_{1'=1^+}, \qquad (3.1)$$

where the three-point function $\chi(1,2,1')$ is defined by

$$\chi_{ij}=\langle T[\psi_i(1)\mathbf{u}(2)\psi_j^\dagger(1')]\rangle. \qquad (3.2)$$

[In writing (3.1) we have again neglected, for the same reason as given below Eq. (2.6), terms arising from the nonzero commutator of the time derivative and time-ordering operations.] Because of the homogeneity of space and time χ depends only on two space-time coordinate differences, so that $\chi=\chi(1\text{–}2,2\text{–}1')$. Since, furthermore, this quantity has periodicity properties along the imaginary time axis, one may make Fourier integral and series expansions in both space-time arguments. We shall denote the Fourier transform by $\chi(\mathbf{p}+\mathbf{q},\zeta_l+\nu_m;\mathbf{p},\zeta_l)$. Inserting the Fourier expansion for χ into (3.1) one obtains the following expression for $P(q,\nu_m)$:

$$P(q,\nu_m)=-\frac{1}{2m}\frac{i}{\beta}\sum_{\zeta_l}e^{\zeta_l0^+}\int\frac{d^3p}{(2\pi)^3}[(\zeta_l+\nu_m)\mathbf{p}+\zeta_l(\mathbf{p}+\mathbf{q})]\cdot\mathrm{Tr}\{\tau_3\chi(\mathbf{p}+\mathbf{q},\zeta_l+\nu_m;\mathbf{p},\zeta_l)\}. \qquad (3.3)$$

In Appendix B we have derived the following double spectral representation for χ (see also Ref. 12)

$$\chi(\mathbf{p}+\mathbf{q},\zeta_l+\nu_m;\mathbf{p},\zeta_l)=\int_{-\infty}^\infty\int_{-\infty}^\infty\frac{d\omega_1d\omega_2}{(2\pi)^2}\left[\frac{\mathbf{f}^{(1)}(\mathbf{p}+\mathbf{q},\omega_1;\mathbf{p},\omega_2)}{(\zeta_l-\omega_1)(\nu_m-\omega_2)}+\frac{\mathbf{f}^{(2)}(\mathbf{p}+\mathbf{q},\omega_1;\mathbf{p},\omega_2)}{(\zeta_l+\nu_m-\omega_1)(\nu_m-\omega_2)}\right]. \qquad (3.4)$$

The explicit forms of the spectral functions $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ are given in Appendix B. We insert Eq. (3.4) into Eq. (3.3), perform the sum over ζ_l, and introduce the result for $P(q,\nu_m)$ into (2.5). The result for K in terms of $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ is then found to be

$$K=\frac{1}{12mT}\int\frac{d^3pd\omega}{(2\pi)^4}\mathbf{p}\cdot\mathrm{Tr}\left\{\tfrac{1}{2}\beta\omega\,\mathrm{sech}^2(\tfrac{1}{2}\beta\omega)\tau_3\mathbf{f}_\mathbf{p}^{(2)}(\omega,0)-(e^{\beta\omega}+1)^{-1}\left[\tau_3(\mathbf{f}_\mathbf{p}^{(1)}(\omega,0)+\mathbf{f}_\mathbf{p}^{(2)}(\omega,0))+2\omega\tau_3\frac{\partial}{\partial\nu}(\mathbf{f}_\mathbf{p}^{(1)}(\omega,\nu)+\mathbf{f}_\mathbf{p}^{(2)}(\omega+\nu,\nu))|_{\nu=0}\right]\right\}. \qquad (3.5)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Here we have used the abbreviation $\mathbf{f}_p^{(i)}(\omega,\nu)=\mathbf{f}^{(i)}(\mathbf{p},\omega,\mathbf{p},\nu)$, $i=1, 2$.

To proceed further we must determine the spectral functions $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ in some approximation. We consider the integral equation for χ in the ladder approximation which is represented by the diagrams in Fig. 2 and is given by

$$\chi(\mathbf{p},\zeta_l+\nu_m;\mathbf{p},\zeta_l)=-\frac{1}{2m}\mathbf{p}(2\zeta_l+\nu_m)\mathcal{G}(p,\zeta_l+\nu_m)\tau_3\mathcal{G}(p,\zeta_l)+\frac{1}{2m}\mathbf{p}[\mathcal{G}(p,\zeta_l+\nu_m)\tau_3+\tau_3\mathcal{G}(p,\zeta_l)]$$
$$-\frac{1}{\beta}\sum_{\mu_n}\int\frac{d^3k}{(2\pi)^3}\mathcal{G}(p,\zeta_l+\nu_m)\tau_3\chi(\mathbf{p}-\mathbf{k},\zeta_l+\nu_m-\mu_n;\mathbf{p}-\mathbf{k},\zeta_l-\mu_n)\tau_3\mathcal{G}(p,\zeta_l)D(\mathbf{k},\mu_n). \quad (3.6)$$

Here $D(\mathbf{k},\mu_n)$ is the phonon propagator ($\mu_n=2\pi ni/\beta$ with n an integer) which has the spectral representation

$$D(\mathbf{k},\mu_n)=\int_{-\infty}^{\infty}\frac{d\nu}{2\pi}\frac{d(k,\nu)}{\mu_n-\nu}. \quad (3.7)$$

Introducing this spectral representation as well as those for $\mathcal{G}$ and χ into (3.6) and doing the sum over n we find

$$\int\frac{d\omega_1 d\omega_2}{(2\pi)^2}\left[\frac{\mathbf{f}_p^{(1)}(\omega_1,\omega_2)}{(\zeta_l-\omega_1)(\nu_m-\omega_2)}+\frac{\mathbf{f}_p^{(2)}(\omega_1,\omega_2)}{(\zeta_l+\nu_m-\omega_1)(\nu_m-\omega_2)}\right]$$
$$=-\frac{\mathbf{p}}{2m}(2\zeta_l+\nu_m)\int\frac{d\omega_1 d\omega_2}{(2\pi)^2}\frac{a(p,\omega_1)\tau_3 a(p,\omega_2)}{(\zeta_l+\nu_m-\omega_1)(\zeta_l-\omega_2)}$$
$$+\frac{\mathbf{p}}{2m}\left[\int\frac{d\omega_1}{2\pi}\left(\frac{a(p,\omega_1)\tau_3}{\zeta_l+\nu_m-\omega_1}+\frac{\tau_3 a(p,\omega_1)}{\zeta_l-\omega_1}\right)\right]+\int\frac{d^3k}{(2\pi)^3}\int\frac{d\nu_1 d\nu_2 d\nu_3 d\mu_1 d\mu_2}{(2\pi)^5}[f(-\mu_1)+N(\nu_3)]d(k,\nu_3)$$
$$\times\frac{a(p,\nu_1)\tau_3}{\zeta_l+\nu_m-\nu_1}\left[\frac{\mathbf{f}_{p-k}^{(1)}(\mu_1,\mu_2)}{\zeta_l-\mu_1-\nu_3}+\frac{\mathbf{f}_{p-k}^{(2)}(\mu_1,\mu_2)}{\zeta_l+\nu_m-\mu_1-\nu_3}\right]\frac{\tau_3 a(p,\nu_2)}{(\nu_m-\mu_2)(\zeta_l-\nu_2)}. \quad (3.8)$$

The Fermi and Planck functions are abbreviated by f and N, respectively. From Eq. (3.8) we can derive two coupled equations for $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ by calculating first the discontinuity across the real ζ_l axis at $\zeta_l=\omega$ for fixed imaginary ν_m and then the discontinuity across the real ν_m axis at $\nu_m=\nu$, or by carrying out these operations in the opposite order. The two equations are

$$\mathbf{f}_p^{(1)}(\omega,\nu)=-(\mathbf{p}/2m)(2\omega+\nu)a(p,\omega+\nu)\tau_3 a(p,\omega)+(2\pi)^{-8}\int d^3k d\nu_1 d\nu_2 d\nu_3 d\mu_1 d\mu_2[f(-\mu_1)+N(\nu_3)]d(k,\nu_3)$$
$$\times a(p,\nu_1)\tau_3\{\mathbf{f}_{p-k}^{(1)}(\mu_1,\mu_2)2\,\mathrm{Im}[(\omega-\nu_3-\mu_1-i0^+)^{-1}(\omega-\nu_2-i0^+)^{-1}]$$
$$\times 2\,\mathrm{Im}[(\omega+\nu-\nu_1-i0^+)^{-1}(\nu-\mu_2-i0^+)^{-1}]+\mathbf{f}_{p-k}^{(2)}(\mu_1,\mu_2)2\pi\delta(\omega-\nu_2)$$
$$\times 2\,\mathrm{Im}[(\nu-\mu_2-i0^+)^{-1}(\omega+\nu-\nu_1-i0^+)^{-1}(\omega+\nu-\nu_3-\mu_1-i0^+)^{-1}]\}\tau_3 a(p,\nu_2), \quad (3.9)$$

and

$$\mathbf{f}_p^{(1)}(\omega,\nu)+\mathbf{f}_p^{(2)}(\omega+\nu,\nu)=(2\pi)^{-7}\int d^3k d\nu_1 d\nu_2 d\nu_3 d\mu_1[f(-\mu_1)+N(\nu_3)]d(k,\nu_3)a(p,\nu_1)\tau_3$$
$$\times\{\mathbf{f}_{p-k}^{(1)}(\mu_1,\nu)2\,\mathrm{Im}[(\omega+\nu-\nu_1-i0^+)^{-1}(\omega-\nu_3-\mu_1-i0^+)^{-1}(\omega-\nu_2-i0^+)^{-1}]+\mathbf{f}_{p-k}^{(2)}(\mu_1,\nu)$$
$$\times 2\,\mathrm{Im}[(\omega+\nu-\nu_3-\mu_1-i0^+)^{-1}(\omega+\nu-\nu_1-i0^+)^{-1}(\omega-\nu_2-i0^+)^{-1}]\}\tau_3 a(p,\nu_2). \quad (3.10)$$

We note that the second and third terms on the right of Eqs. (3.6) and (3.8), which are discontinuity terms of the type we have previously omitted, do not in fact contribute to the double spectral functions of physical interest. Let us now consider Eq. (3.10) in the limit $\nu\to 0$. We make the Ansatz

$$\mathbf{f}_p^{(i)}(\omega,0)=(\mathbf{p}/m)g^{(i)}(\epsilon_p,\omega), \quad i=1, 2, \quad (3.11)$$

and obtain from (3.10) an equation for $g^{(1)}+g^{(2)}$. On the right-hand side of this equation there occurs the angular factor $(\mathbf{p}-\mathbf{k})\cdot\mathbf{p}/p^2$, which may be replaced by 1 if one neglects terms of order k_BT/μ and k_BT/cp_F (where c is a typical velocity of sound). The d^3k integration may now be transformed into integrations over k, or the phonon

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

frequencies $\omega_\lambda(k)$, where λ designates the phonon branch, and over ϵ_{p-k}. Since the latter integral runs over practically all ϵ_{p-k}, we introduce the functions

$$g^{(i)}(\omega)=\int_{-\infty}^{\infty} d\epsilon g^{(i)}(\epsilon,\omega) \tag{3.12}$$

and integrate both sides of the equation for $g^{(1)}(\epsilon,\omega)+g^{(2)}(\epsilon,\omega)$ over all ϵ. The result can be written in the form

$$g^{(1)}(\omega)+g^{(2)}(\omega)=(2\pi)^{-4}\int d\epsilon d\nu_1 d\nu_2 d\nu_3 d\mu_1 \sum_\lambda \int d\omega_\lambda F_\lambda(\omega_\lambda) d_\lambda(\omega_\lambda,\nu_3)[f(-\mu_1)+N(\nu_3)]a(\epsilon,\nu_1)$$
$$\times\tau_3[g^{(1)}(\mu_1)+g^{(2)}(\mu_1)]\tau_3 a(\epsilon,\nu_2)2\ \mathrm{Im}[(\omega-\nu_1-i0^+)^{-1}(\omega-\nu_3-\mu_1-i0^+)^{-1}(\omega-\nu_2-i0^+)^{-1}]. \tag{3.13}$$

Above we have followed Schrieffer *et al.* in characterizing the phonons by frequency distributions $F_\lambda(\omega_\lambda)$ and spectral functions $d_\lambda(\omega_\lambda,\nu)$ which are of the form

$$d_\lambda(\omega_\lambda,\nu)=2\pi\alpha_\lambda^2(\omega_\lambda)[\delta(\nu-\omega_\lambda)-\delta(\nu+\omega_\lambda)], \tag{3.14}$$

where α_λ measures the strength of the electron-phonon coupling. All of the constants arising from transforming the variables of integration are contained in α_λ.

The functions $g^{(i)}(\omega)$ are matrices in the Nambu space. We make the decomposition

$$g^{(i)}(\omega)=g_0^{(i)}(\omega)1+g_1^{(i)}(\omega)\tau_1 +g_2^{(i)}(\omega)\tau_2+g_3^{(i)}(\omega)\tau_3. \tag{3.15}$$

On the left side of (3.13) we are interested only in the τ_3 component because only this enters the integral for the thermal conductivity. One finds that the terms containing $g_2^{(i)}$ on the right of (3.13) do not enter the τ_3 component of the equation. The terms proportional to $g_0^{(i)}$ and $g_1^{(i)}$ make no contribution to the τ_3 component when integrated over ϵ since the integrands are odd in ϵ. We are then left with a homogeneous integral equation for $g_3^{(1)}(\omega)+g_3^{(2)}(\omega)$. We may now explicitly do the ϵ integral and perform the sum in the Nambu space. When this is done one finds that the right-hand side is *zero* because there occurs an integral we have met with in Sec. II, namely,

$$I_1=\int_{-\infty}^{\infty} d\epsilon\ \mathrm{Tr}[\tau_3 \mathcal{G}(\epsilon,\omega-i0^+)\tau_3\mathcal{G}(\epsilon,\omega-i0^+)]=0. \tag{3.16}$$

(See Appendix C for the evaluation.)

We have thus shown that

$$\frac{1}{2\mu}\int d\epsilon\mathbf{p}\cdot\mathrm{Tr}\{\tau_3[\mathbf{f}_p^{(1)}(\omega,0)+\mathbf{f}_p^{(2)}(\omega,0)]\} =g_3^{(1)}(\omega)+g_3^{(2)}(\omega)=0. \tag{3.17}$$

Hereafter we shall write

$$g_3^{(1)}(\omega)=-g_3^{(2)}(\omega)\equiv g(\omega). \tag{3.18}$$

One can also show, using time reversal invariance and (3.17), that in our approximation we may take $g(\omega)$ to be real (Appendix B).

Using (3.17) in the general expression for the thermal conductivity (3.5) we see that the second term in this last equation does not contribute. The third term on the right-hand side of (3.5) also does not contribute. To see this we take the derivative of both sides of Eq. (3.10) with respect to ν and then set $\nu=0$. We now remove the vector dependence by an Ansatz similar to (3.11) and integrate the resulting equation for scalar quantities over all ϵ. The procedure is identical to that described above Eq. (3.13). We then take the τ_3 projection and find that the right-hand side is zero either because of oddness in ϵ or because of (3.16) and (3.17). In other words, in our approximation we have

$$\int d\epsilon\mathbf{p}\cdot\mathrm{Tr}\left\{\tau_3\frac{\partial}{\partial\nu}[\mathbf{f}_p^{(1)}(\omega,\nu)+\mathbf{f}_p^{(2)}(\omega+\nu,\nu)]|_{\nu=0}\right\}=0. \tag{3.19}$$

We therefore see that the thermal conductivity (3.5) is entirely determined by the τ_3 component of $\mathbf{p}\cdot\mathbf{f}_p^{(2)}(\omega,0)$ or equivalently, through Eqs. (3.11), (3.12), (3.15), and (3.18), by the scalar function $g(\omega)$. To determine this function let us consider Eq. (3.9) in the limit $\nu=0$. In the curly bracket on the right of (3.9) there are eight terms which arise from explicitly taking the imaginary parts. It is convenient to increase this number from eight to ten by adding and subtracting the term $4\pi^4\mathbf{f}_{p-k}^{(1)}(\mu_1,\mu_2)\delta(\omega-\nu_3-\mu_1)\delta(\omega-\nu_2)\delta(\omega-\nu_1)\delta(\mu_2)$ in the curly bracket. The resulting equation for $\mathbf{f}_p^{(1)}(\omega,0)$ may be written in the form

$$\mathbf{f}_p^{(1)}(\omega,0)=-a(p,\omega)\tau_3\mathbf{\Gamma}(\mathbf{p},\omega)\tau_3 a(p,\omega)$$
$$+(2\pi)^{-4}\left\{\int d^3k d\nu[f(\nu-\omega)+N(\nu)]d(k,\nu)\tfrac{1}{2}[\mathcal{G}(p,\omega-i0^+)\tau_3\mathbf{f}_{p-k}^{(1)}(\omega-\nu,0)\tau_3\mathcal{G}(p,\omega+i0^+)\right.$$
$$\left.+\mathcal{G}(p,\omega+i0^+)\tau_3\mathbf{f}_{p-k}^{(1)}(\omega-\nu,0)\tau_3\mathcal{G}(p,\omega-i0^+)]\right\}+\mathbf{f}_p'(\omega,0). \tag{3.20}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(The occurrence of the $+i0^+$'s in the arguments of the Green's functions in the curly bracket will be important in what follows.) In (3.20) $\mathbf{\Gamma}(\mathbf{p},\omega)$ describes the corrected vertex associated with the ladder approximation and is given by

$$\mathbf{\Gamma}(\mathbf{p},\omega)=\omega(\mathbf{p}/m)\tau_3-\int\frac{d^3k}{(2\pi)^3}\int\frac{d\mu_1 d\mu_2 d\nu_3}{(2\pi)^3}\{[f(-\mu_1)+N(\nu_3)]d(k,\nu_3)[\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(1)}(\mu_1,\mu_2)+\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(2)}(\mu_1,\mu_2)]$$
$$\times\mathrm{Re}[(\omega-\nu_3-\mu_1-i0^+)^{-1}(-\mu_2-i0^+)^{-1}]\}. \quad (3.21)$$

The function $\mathbf{f}_p'(\omega,0)$ consists of four interference terms. The explicit form is

$$\mathbf{f}_p'(\omega,0)=2\int\frac{d^3k}{(2\pi)^3}\int\frac{d\mu_1 d\nu_3}{(2\pi)^2}[f(-\mu_1)+N(\nu_3)]d(k,\nu_3)P\frac{1}{\omega-\nu_3-\mu_1}[\mathrm{Re}\mathcal{G}(p,\omega-i0^+)]\tau_3[\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(1)}(\mu_1,0)+\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(2)}(\mu_1,0)]$$
$$\times\tau_3[\mathrm{Im}\mathcal{G}(p,\omega-i0^+)]-2\int\frac{d^3k}{(2\pi)^3}\int\frac{d\mu_2 d\nu_3}{(2\pi)^2}[f(\nu_3-\omega)+N(\nu_3)]d(k,\nu_3)P\left(\frac{1}{\mu_2}\right)$$
$$\times\{[\mathrm{Im}\mathcal{G}(p,\omega-i0^+)]\tau_3\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(1)}(\omega-\nu_3,\mu_2)\tau_3[\mathrm{Re}\mathcal{G}(p,\omega-i0^+)]$$
$$+[\mathrm{Re}\mathcal{G}(p,\omega-i0^+)]\tau_3\mathbf{f}_{\mathbf{p}-\mathbf{k}}{}^{(2)}(\omega-\nu_3,\mu_2)\tau_3[\mathrm{Im}\mathcal{G}(p,\omega-i0^+)]\}. \quad (3.22)$$

We shall return in a moment to $\mathbf{f}'$. However, let us first consider the vertex corrections (3.21). In general, we may define the Fourier coefficients of a vertex operator by the relation

$$\chi(\mathbf{p},\zeta_l+\nu_m;\mathbf{p},\zeta_l)$$
$$=(\mathbf{p}/2m)[\mathcal{G}(p,\zeta_l+\nu_m)\tau_3+\tau_3\mathcal{G}(p,\zeta_l)]$$
$$-\mathcal{G}(p,\zeta_l+\nu_m)\mathbf{\Gamma}_p(\zeta_l+\nu_m,\zeta_l)\mathcal{G}(p,\zeta_l). \quad (3.23)$$

The function $\mathbf{\Gamma}(\mathbf{p},\omega)$ is then related to the analytic continuation of the Fourier series coefficients by

$$\mathbf{\Gamma}(p,\omega)=\mathbf{\Gamma}_p(\omega+\nu,\nu)|_{\nu=0}, \quad (3.24)$$

where

$$\mathbf{\Gamma}_p(\omega+\nu,\nu)=\tfrac{1}{2}\{\mathbf{\Gamma}_p(\zeta_l+\nu_m,\zeta_l)|_{\substack{\zeta_l=\omega-i0^+\\ \nu_m=\nu-i0^+}}$$
$$+\mathbf{\Gamma}_p(\zeta_l+\nu_m,\zeta_l)|_{\substack{\zeta_l=\omega+i0^+\\ \nu_m=\nu+i0^+}}\}. \quad (3.25)$$

It is straightforward to verify that $\mathbf{\Gamma}(\mathbf{p},\omega)$ as defined by (3.23)–(3.25) is in fact given by (3.21) in the ladder approximation of (3.6).

There exists a Ward identity, reflecting the conservation of energy density, for the vertex function introduced above. This Ward identity has been discussed with great clarity by Langer,[8] and it is therefore not necessary to give the derivation here. The result in terms of Fourier series coefficients is[16]

$$\mathbf{\Gamma}_p(\zeta_l,\zeta_l)=-\zeta_l\nabla_p\mathcal{G}^{-1}(p,\zeta_l)$$
$$=-\zeta_l\nabla_p[\zeta_l-((p^2/2m)-\mu)\tau_3-\Sigma(p,\zeta_l)]. \quad (3.26)$$

[16] In Langer's result (Ref. 8) there occur additional terms due to the nonlocality in space of the Coulomb interaction he considers. Such terms do not occur for the electron-phonon interaction. For the direct Coulomb interaction it is in any case consistent with the Eliashberg approximation to neglect such terms. It should also be remarked that since the Ward identity (3.26) follows from the strict conservation of energy density it requires the inclusion in the vertex function of the heat current $\mathbf{u}_{ph}$ introduced above Eq. (2.2) and then rapidly discarded. This means that in using (3.26) we are including a part of the "phonon-drag" thermal conduction. Since we have already argued that processes of this type have a negligible effect on the thermal conductivity, their inclusion here is of no consequence.

We notice that in the Eliashberg approximation of neglecting the momentum dependence of the self-energy operator the vertex correction is zero. We may therefore make the replacement

$$\mathbf{\Gamma}(\mathbf{p},\omega)\rightarrow(\mathbf{p}/m)\omega\tau_3. \quad (3.27)$$

In a more exact treatment one would have to include the effective mass and wave function renormalization corrections implied by (3.26).

We now return to Eq. (3.20), remove the vector dependence, and integrate both sides of the resulting equation over all ϵ. The τ_3 component of this equation is then found to be

$$g(\omega)=-\frac{\pi\omega}{|\mathrm{Im}Z(\omega^2-\Delta^2)^{1/2}|}\left(1+\frac{\omega^2-|\Delta|^2}{|\omega^2-\Delta^2|}\right)$$
$$\times\left\{1-\frac{1}{2\omega}\int\frac{d\nu}{2\pi}\sum_\lambda\int d\omega_\lambda F_\lambda(\omega_\lambda)d_\lambda(\omega_\lambda,\nu)\right.$$
$$\left.\times[f(\nu-\omega)+N(\nu)]g(\omega-\nu)\right\}+g'(\omega), \quad (3.28)$$

where $g'(\omega)$ comes from the interference terms $\mathbf{f}'$ given in (3.22). Whereas in the other terms of (3.20) the τ_3 component on the left is only coupled to the τ_3 component on the right, g' turns out to depend only on g_2. The

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

explicit form we find is

$$g'(\omega)=\frac{i\pi\omega\,\mathrm{Im}\Delta}{|\mathrm{Im}Z(\omega^2-\Delta^2)^{1/2}||\omega^2-\Delta^2|}$$
$$\times\Big\{\int\frac{d\nu d\mu_1}{(2\pi)^2}[f(-\mu_1)+N(\nu)]$$
$$\times\sum_\lambda\int d\omega_\lambda\frac{F_\lambda(\omega_\lambda)d_\lambda(\omega_\lambda,\nu)}{(\omega-\nu-\mu_1)_p}[g_2^{(1)}(\mu_1)+g_2^{(2)}(\mu_2)]$$
$$+\int\frac{d\nu d\mu_2}{(2\pi)^2}[f(\nu-\omega)+N(\nu)]$$
$$\times\sum_\lambda\int d\omega_\lambda\frac{F_\lambda(\omega_\lambda)d_\lambda(\omega_\lambda,\nu)}{(\mu_2)_p}$$
$$\times[g_2^{(1)}(\omega-\nu,\mu_2)-g_2^{(2)}(\omega-\nu,\mu_2)]\Big\}. \quad (3.29)$$

The coupling between the τ_3 and the τ_2 components, though physically interesting, appears to be of negligible importance for our problem. In the first place, an examination of Eq. (3.13) shows that the τ_2 component on the left is coupled only to the τ_2 component on the right. In other words, $g_2^{(1)}(\omega)+g_2^{(2)}(\omega)$ satisfies a linear homogeneous integral equation. Therefore this quantity can at most be nonzero for certain values of the electron-phonon coupling constant, an unphysical possibility which it seems safe to ignore. In comparing the remaining term in Eq. (3.29) with the homogeneous term in (3.28) we encounter the ratio $[\mathrm{Im}\Delta(\omega)/\omega]$. From the numerical results of Schrieffer *et al.* we see that for the frequencies of interest, i.e., $\omega<10kT_c$, this factor is less than 0.1 at $T=0$ and consequently exceedingly small just below $T=T_c$. Thus, barring the occurrence of other large factors, the second term in (3.29) is also negligible. An order of magnitude comparison of the integrands of this term and the homogeneous term in (3.28) yields the ratio $\Gamma(\omega-\nu)/(\omega-\nu)$, where $\Gamma(\omega-\nu)$ is a typical width at frequency $(\omega-\nu)$. Although this factor is not small over the entire region of interest, it is never large. It therefore seems safe to conclude that the terms in g' are of no significance. These terms evidently describe weak-relaxation processes peculiar to the superconducting state which involve the motion of condensed pairs relative to the single-particle excitations.[17] We plan to study these processes further.

Since the coupling between different components in the Nambu space came solely from g' in (3.28), none remains when this term is omitted. On introducing the spectral function (3.14) we then obtain

$$g(\omega)=-\frac{\pi\omega}{|\mathrm{Im}Z(\omega^2-\Delta^2)^{1/2}|}\left(1+\frac{\omega^2-|\Delta|^2}{|\omega^2-\Delta^2|}\right)$$
$$\times\Big\{1-(2\omega)^{-1}f^{-1}(-\omega)\sum_\lambda\int d\omega_\lambda\alpha_\lambda^2(\omega_\lambda)F_\lambda(\omega_\lambda)$$
$$\times[f(\omega_\lambda-\omega)(1+N(\omega_\lambda))g(\omega-\omega_\lambda)$$
$$+f(-\omega_\lambda-\omega)N(\omega_\lambda)g(\omega+\omega_\lambda)]\Big\}. \quad (3.30)$$

It follows from (3.30) that $g(\omega)$ is odd in ω, which one also expects on general grounds. Further one sees that near $T=0$ $g(\omega)=0$ for $\omega<\Delta_0$. From (3.5) and our subsequent analysis we see that the thermal conductivity K is determined once $g(\omega)$ is known by the quadrature

$$K=-\frac{n}{4mk_BT^2}\int_0^\infty\frac{d\omega}{2\pi}\omega\,\mathrm{sech}^2(\tfrac{1}{2}\beta\omega)g(\omega). \quad (3.31)$$

The Hartree-Fock result (2.17) is recovered if we approximate $g(\omega)$ by the inhomogeneous term of (3.30). In the small width limit $|\mathrm{Im}[Z(\omega^2-\Delta^2)^{1/2}]|^{-1}\to(\omega^2-\Delta_1^2)^{1/2}/\omega\Gamma$. If one introduces instead of $g(\omega)$ a new function $G(\omega)$ defined according to

$$g(\omega)=-\pi\,\mathrm{sgn}\omega(\omega^2-\Delta_1^2)^{1/2}\left(1+\frac{\omega^2-\Delta_1^2}{|\omega^2-\Delta_1^2|}\right)G(\omega), \quad (3.32)$$

one obtains an integral equation for $G(\omega)$ which is identical to the integral equation obtained previously from the Boltzmann equation. [Essentially Eq. (2.18) of I if the differences in phonon spectra used are taken into account.]

If the width is not small, (3.30) is the most convenient form for numerical calculations. The complex functions $\Delta(\omega)$ and $Z(\omega)$ have to be determined from the set of integral equations given in Appendix A.

IV. CONCLUSIONS

In spite of some anfractuous analysis, we have been led to a relatively straightforward scheme for calculating the electronic thermal conductivity of strongly coupled superconductors. The Eliashberg approximation of weak momentum dependence is essential for much of the final simplicity. We have found in this approximation that the thermal conductivity is determined by an integral of the expression $\omega\,\mathrm{sech}^2(\frac{1}{2}\beta\omega)g(\omega)$ over all excitation energies ω. The function $g(\omega)$ plays the role of the deviation from an equlibrium distribution and satisfies the generalized Boltzmann equation (3.30).

If one neglects the difference between the transport and the scattering cross sections, or equivalently keeps only the inhomogeneous term of the integral equation (3.30), one recovers the result of the Hartree-Fock

[17] This is a situation reminiscent of the collective excitation which occurs in the response of the superconductor to finite frequency and wave number electromagnetic disturbances.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

approximation of Sec. II. In this case, for frequencies not too far above the gap edge $g(\omega)$ becomes approximately equal to $-2\pi(\omega^2-\Delta_1^2)^{1/2}/\Gamma(\omega)$, where $\Gamma(\omega)$ plays the role of a width parameter and is in the weak-coupling limit identical with the reciprocal of twice the quasiparticle lifetime as calculated in II. The expression for $\Gamma(\omega)$ is $\Gamma(\omega)\approx|Z_2(\omega^2-\Delta_1^2)-Z_1\Delta_1\Delta_2|\omega^{-1}$. The renormalization function $Z(\omega)$ and the gap function $\Delta(\omega)$ have to be determined from the Eliashberg gap equations given in Appendix A.

The gap equations have been solved for lead by Schrieffer *et al.* in the limit of zero temperature, but not for temperatures near T_c which are of interest here. We can therefore only speculate about the manner in which the great reduction in the thermal conductivity of superconducting lead might be explained by our theory. The calculations in I have shown that in the small width limit the decrease of K_s/K_n with decreasing temperature is mainly due to the increase of Δ_1 in the factor $(\omega^2-\Delta_1^2)^{1/2}$ in the integrand of K_s, while $T^3\Gamma_s^{-1}$ (as a function of $\beta\omega$) is relatively insensitive to changes in temperature near T_c. Let us assume that this is also true for lead. Then it is important to note that near T_c energies up to about $\omega_m=10k_BT_c$ contribute to the heat current. This means that the first peak in the curve for $\Delta_1(\omega)$ calculated by Schrieffer *et al.* which occurs at $\omega_1{}'=4.4\times10^{-3}$ eV$\approx10k_BT_c$ occurs roughly at ω_m. If the shape of the $\Delta_1(\omega)$ curve were preserved at temperatures near T_c or, more precisely, if the ratio $\Delta_1(\omega)/\Delta_0$ were nearly independent of T, then the effective gap Δ_1 for the excitations predominantly carrying the heat current would clearly be much greater than the gap edge Δ_0. This would lead to a substantial reduction of K_s/K_n compared to the value obtained in the small width limit.

To go beyond the Hartree-Fock approximation one must numerically solve the integral equation (3.30). From the experience of the calculations reported in I, it is not expected that this more exact treatment will drastically affect the thermal conductivity. Nevertheless, without a solution of the integral equation it seems difficult to estimate the errors introduced by the Hartree-Fock approximation. It is worth noting that the most complicated part of the integral equation (3.30), which depends on the solution of the gap equations, is the Hartree-Fock value of $g(\omega)$ which occurs as a factor in both the homogeneous and inhomogeneous terms. It should also be noted that the same electron-phonon coupling strengths, α_λ, and phonon frequency distributions $F_\lambda(\omega_\lambda)$ occur in the integral equation for $g(\omega)$ and those for $\Delta(\omega)$ and $Z(\omega)$.

In the small width limit, i.e., $\alpha_\lambda\to0$, we have indicated how (3.30) may be transformed into the Bloch-Boltzmann equation for the superconducting state. A byproduct of the analysis given in the present paper has thus been the justification of the use of this equation, as in I, for example. In our extreme long-wavelength low-frequency limit the only validity criterion for the use of the Boltzmann equation is that well defined quasiparticles exist in the thermally excited frequency range below roughly $10k_BT$. This criterion also applies to the normal state, and when it fails one must resort to the more general approach described in this paper.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor L. P. Kadanoff for a useful discussion. Much of the contribution of the first named author (V. A.) was performed during a six week visit to the Bell Telephone Laboratories during the summer of 1963. The authors are happy to take the opportunity to thank the Bell Telephone Laboratories and in particular their friends and colleagues in the Theoretical Physics Department for their gracious hospitality.

APPENDIX A

In the Nambu representation the superconductive approximation to the electronic self-energy operator has the Fourier coefficients

$$\Sigma(p,\zeta_l)=-\frac{1}{\beta}\sum_{\nu_m}\int\frac{d^3q}{(2\pi)^3}\tau_3 \times\mathcal{G}(p-q,\zeta_l-\nu_m)\tau_3D(q,\nu_m). \quad \text{(A1)}$$

Here

$$\mathcal{G}^{-1}(p,z)=z-\epsilon_p-\Sigma(p,z) \quad \text{(A2)}$$

and the Nambu Ansatz for $\Sigma(p,z)$ is

$$\Sigma(p,z)=z\zeta_p(z)+\chi_p(z)\tau_3+\phi_p(z)\tau_1. \quad \text{(A3)}$$

Following Eliashberg we neglect χ the even in frequency component of the diagonal (in Nambu space) part of Σ as contributing only a shift in the chemical potential. Then $\mathcal{G}^{-1}$ has the form given in Eq. (2.12), namely

$$\mathcal{G}^{-1}(p,z)=zZ(z)-\epsilon\tau_3-\phi(z)\tau_1, \quad \text{(A4)}$$

where $Z(z)=1-\zeta(z)$. The functions ϕ and Z have the properties

$$\begin{aligned} &Z(z^*)=Z^*(z);\quad Z(-z)=Z(z),\\ &\phi(z^*)=\phi^*(z);\quad \phi(-z)=\phi(z). \end{aligned} \quad \text{(A5)}$$

$Z(z)$ is even because $z\zeta(z)$ is the odd component of the diagonal part of Σ. The function $\phi(z)$ may be taken to be even in z for s state pairing because of spin-rotation invariance, time-reversal symmetry, and spatial-inversion symmetry.[18] Using (A5) it is easily seen that the spectral function $a(p,\omega)$ defined by Eq. (2.13) has the properties

$$\begin{aligned} &\mathrm{Tr}[a(p,\omega)]=\mathrm{Tr}[a(p,-\omega)],\\ &\mathrm{Tr}[\tau_i a(p,\omega)]=-\mathrm{Tr}[\tau_i a(p,-\omega)]. \end{aligned} \quad \text{(A6)}$$

Now we insert the spectral representations for $\mathcal{G}$ and D

[18] Questions of symmetry are very carefully considered in P. Nozieres, *Theory of Interacting Fermion Systems* (W. A. Benjamin, Inc., New York, 1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

[Eqs. (2.10) and (3.7)] into (A1) and perform the sum over ν_m. We then make the weak-momentum approximation for ϕ and Z and perform the integral over d^3q in the manner discussed in Sec. III. The integral of $a(\epsilon,\omega)$ over all ϵ which is needed here is given by[6,10]

$$\int_{-\infty}^{\infty} d\epsilon a(\epsilon,\omega) = -2\pi \,\mathrm{Re}\left\{\frac{\omega Z(\omega)+\phi(\omega)\tau_1}{\epsilon_1(\omega)}\right\}, \quad \text{(A7)}$$

where ϵ_1 is the root of $\epsilon^2+\phi^2(\omega)-(\omega-i0^+)^2Z^2(\omega)=0$ which lies in the upper half-plane. [For real argument ω, $\phi(\omega)$ and $Z(\omega)$ are given by (2.16).] After performing the steps outlined above, we decompose the integral equation for $\Sigma(z)$ in the Nambu space and equate coefficients of 1 and τ_3 on the right and left sides. The symmetries (A6) may be used to simplify the equations. The final equations are

$$\Delta(z)=\frac{1}{Z(z)}\int_0^\infty d\omega_1 \,\mathrm{Re}\left\{\frac{\Delta(\omega_1)}{[\omega_1^2-\Delta^2(\omega_1)]^{1/2}}\right\}\sum_\lambda \int d\omega_\lambda \alpha_\lambda^2(\omega_\lambda)F_\lambda(\omega_\lambda)$$
$$\times\{[N(\omega_\lambda)+f(-\omega_1)][(\omega_1+\omega_\lambda+z)^{-1}+(\omega_1+\omega_\lambda-z)^{-1}]$$
$$-[N(\omega_\lambda)+f(\omega_1)][(-\omega_1+\omega_\lambda+z)^{-1}+(-\omega_1+\omega_\lambda-z)^{-1}]\}, \quad \text{(A8)}$$

$$z[1-Z(z)]=\int_0^\infty d\omega_1 \,\mathrm{Re}\left\{\frac{\omega_1}{[\omega_1^2-\Delta^2(\omega_1)]^{1/2}}\right\}\sum_\lambda \int d\omega_\lambda \alpha_\lambda^2(\omega_\lambda)F_\lambda(\omega_\lambda)$$
$$\times\{[N(\omega_\lambda)+f(-\omega_1)][(\omega_1+\omega_\lambda+z)^{-1}-(\omega_1+\omega_\lambda-z)^{-1}]$$
$$+[N(\omega_\lambda)+f(\omega_1)][(-\omega_1+\omega_\lambda+z)^{-1}-(-\omega_1+\omega_\lambda-z)^{-1}]\}.$$

The physical values of Δ and Z are obtained by setting $z=\omega-i0^+$ in (A8). The sign of the square root in (A8) is to be determined by $\mathrm{Im}\{Z(\omega^2-\Delta^2)^{1/2}\}<0$.

APPENDIX B

The spectral representation (3.4) for the object $\chi_{ij}(t_1,t_2,t_3)\equiv T\langle\psi_i(t_1)\mathbf{u}(t_2)\psi_j^\dagger(t_3)\rangle$ follows from its periodicity properties for imaginary times. (We suppress the spatial arguments for brevity.) Consider the region of time $\mathrm{Re}t_1,\ t_2,\ t_3=0$; $0>\mathrm{Im}t_1,\ t_2,\ t_3>-\beta\equiv i\tau$, and let T order the operators from right to left in order of increasing negative imaginary time.[7] Then measuring energies with respect to the chemical potential, and using the cyclic invariance of the trace operation contained in the ground canonical ensemble average, one finds the periodicity properties

$$\begin{aligned}\chi_{ij}(0,t_2,t_3)&=-\chi_{ij}(\tau,t_2,t_3),\\ \chi_{ij}(t_1,0,t_3)&=+\chi_{ij}(t_1,\tau,t_3),\\ \chi_{ij}(t_1,t_2,0)&=-\chi_{ij}(t_1,t_2,\tau).\end{aligned} \quad \text{(B1)}$$

As a result we may make the Fourier series expansion

$$\chi_{ij}(t_1,t_2,t_3)=\frac{1}{\tau^3}\sum_{lmn}\chi_{ij}(n,m,l)e^{-i\zeta_n t_1}e^{-i\nu_m t_2}e^{-i\zeta_l t_3}, \quad \text{(B2)}$$

where $\zeta_n=(2n+1)\pi/\tau$, $\nu_m=2\pi m/\tau$, $\zeta_l=(2l+1)\pi/\tau$. Because $\chi(t_1,t_2,t_3)=\chi(t_1-t_2,\ t_2-t_3)$ we expect $\chi(nml)$ to be zero unless $\zeta_n+\nu_m+\zeta_l=0$. Inverting the Fourier series we have

$$\chi_{ij}(n,m,l)=\int_0^\tau dt_1\int_0^\tau dt_2\int_0^\tau dt_3\chi_{ij}(t_1,t_2,t_3)$$
$$\times e^{i\zeta_n t_1}e^{i\nu_m t_2}e^{i\zeta_l t_3}. \quad \text{(B3)}$$

To do the integral we break up the domain of integration into six regions corresponding to the six time-ordered constituents of χ. In each of these regions the time dependence of the Heisenberg operators can be displayed and the integrals trivially performed. One verifies that $\chi(nml)$ is indeed zero unless the condition written above (B3) is satisfied. It is convenient to take as independent variables ζ_l and ν_m and to introduce $\chi(\zeta_l+\nu_m,\zeta_l)\equiv\tau\chi(l+m,-m,-l-1)$ so that the Fourier expansion reads

$$\chi_{ij}(t_1-t_2,\ t_2-t_3)=\frac{1}{\tau^2}\sum_{lm}\chi_{ij}(\zeta_l+\nu_m,\zeta_l)$$
$$\times e^{-i(\zeta_l+\nu_m)(t_1-t_2)}e^{-i\zeta_l(t_2-t_3)}. \quad \text{(B4)}$$

After explicitly carrying out the time integrals in (B3) in the manner indicated above, one finds that the Fourier coefficient $\chi(\zeta_l+\nu_m,\zeta_l)$ can be written in the form (3.4). The spectral functions $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ obtained explicitly in this way are sums of products of three matrix elements, and two delta functions involving the frequencies ω_1 and ω_2. After some rearrangement one finds that the spectral functions may be written compactly in the form

$$\mathbf{f}_{ij}^{(1)}(\omega_1,\omega_2)=\int_{-\infty}^{\infty}dt_1\int_{-\infty}^{\infty}dt_2$$
$$\times\langle\{[\mathbf{u}(-t_2),\psi_i(0)],\psi_j^\dagger(-t_1)\}\rangle$$
$$\times e^{i\omega_1 t_1}e^{i\omega_2 t_2}, \quad \text{(B5)}$$

and

$$\mathbf{f}_{ij}^{(2)}(\omega_1,\omega_2)=\int_{-\infty}^{\infty}dt_1\int_{-\infty}^{\infty}dt_2\langle\{[\mathbf{u}(-t_2),\psi_j^\dagger(0)],\psi_i(t_1)\}\rangle$$
$$\times e^{i\omega_1 t_1}e^{i\omega_2 t_2}. \quad \text{(B6)}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Above, the curly bracket indicates the anticommutator and the square bracket the commutator.

We now wish to substantiate the remarks made after Eq. (3.18) about the reality of $g(\omega)$. If one uses time reversal invariance on the closed forms (B5) and (B6) one finds the relation

$$\mathbf{f}_{ij}^{(1)}(\omega_1,\omega_2)=-(\tau_2)_{ik}\mathbf{f}_{lk}^{(2)}(-\omega_1,\omega_2)^*(\tau_2)_{lj}. \quad \text{(B7)}$$

Further from explicit complex conjugation of (B5) and (B6) one finds

$$\mathbf{f}_{ij}^{(1)}(\omega_1,\omega_2)^*=-\mathbf{f}_{ji}^{(2)}(\omega_1,-\omega_2). \quad \text{(B8)}$$

From (B7) and (B8) we can derive the exact identity

$$\mathrm{Tr}[\mathbf{f}^{(2)}(\omega_1,\omega_2)\tau_3]=-\mathrm{Tr}[f^{(1)}(\omega_1,-\omega_2)\tau_3]^*. \quad \text{(B9)}$$

In the ladder approximation we have proved that

$$\int d\epsilon\mathbf{p}\cdot\mathrm{Tr}\{\tau_3[\mathbf{f}^{(1)}(\omega,0)+\mathbf{f}^{(2)}(\omega,0)]\}=0. \quad \text{(B10)}$$

It follows from (B9) and (B10) that in our approximation

$$\int d\epsilon\mathbf{p}\cdot\mathrm{Tr}[\tau_3\mathbf{f}^{(1)}(\omega,0)] = \int\left\{d\epsilon\mathbf{p}\cdot\mathrm{Tr}[\tau_3\mathbf{f}^{(1)}(\omega,0)]\right\}^*. \quad \text{(B11)}$$

Using the definition of $g(\omega)$ contained in Eqs. (3.18), (3.15), and (3.11) we see that (B11) implies that this function is real.

APPENDIX C

Here we sketch the evaluation of the integrals I_1 and I_2 given in Eqs. (2.14) and (2.15). Although the evaluation is straightforward it seems worth presenting, particularly because (2.14) plays an important role in Sec. III.

Consider first I_1. We have using (2.12)

$$I_1=\int_{-\infty}^{\infty}d\epsilon\,\mathrm{Tr}[\tau_3\mathcal{G}(\epsilon,\omega-i0^+)\tau_3\mathcal{G}(\epsilon,\omega-i0^+)]$$

$$=2\int_{-\infty}^{\infty}d\epsilon\frac{\omega^2Z^2+\epsilon^2-\phi^2}{(\epsilon-\epsilon_1)^2(\epsilon+\epsilon_1)^2}, \quad \text{(C1)}$$

where $\epsilon=\pm\epsilon_1$ are the roots of

$$\omega^2Z^2-\epsilon^2-\phi^2=0. \quad \text{(C2)}$$

We choose $+\epsilon_1$ to lie in the upper half-plane. Then by contour integration

$$I_1=4\pi i\left[-\frac{\omega^2Z^2-\phi^2}{4\epsilon_1^3}+\frac{1}{4\epsilon_1}\right]=0, \quad \text{(C3)}$$

where the last step follows from (C2). The integral I_2, on the other hand, is given by

$$I_2=\int_{-\infty}^{\infty}d\epsilon\,\mathrm{Tr}[\tau_3\mathcal{G}(\epsilon,\omega-i0^+)\tau_3\mathcal{G}(\epsilon,\omega+i0^+)]$$

$$=2\int_{-\infty}^{\infty}d\epsilon\frac{\omega^2|Z|^2+\epsilon^2-|\phi|^2}{(\epsilon-\epsilon_1)(\epsilon+\epsilon_1)(\epsilon-\epsilon_1^*)(\epsilon+\epsilon_1^*)}. \quad \text{(C4)}$$

Since ϵ_1 is in the upper half-plane, so is $-\epsilon_1^*$. By a straightforward contour integration,

$$I_2=4\pi i\left[\frac{\omega^2|Z|^2-|\phi|^2+\epsilon_1^2}{8i\epsilon_1\,\mathrm{Im}\epsilon_1\,\mathrm{Re}\epsilon_1}+\frac{\omega^2|Z|^2-|\phi|^2+(\epsilon_1^*)^2}{8i\,\mathrm{Re}\epsilon_1\,\mathrm{Im}\epsilon_1\cdot\epsilon_1^*}\right]$$

$$=\frac{\pi}{\mathrm{Im}\epsilon_1}\left[\frac{\omega^2|Z|^2-|\phi|^2+|\epsilon_1|^2}{|\epsilon_1|^2}\right]. \quad \text{(C5)}$$

Using (C2) and the relation $\Delta(\omega)=\phi(\omega)/Z(\omega)$, (C5) may be put in the form given in (2.15).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Dissipation in Quantum Mechanics. Two-Level System. II

I. R. SENITZKY
U. S. Army Electronics Research and Development Laboratories, Fort Monmouth, New Jersey

(Received 20 December 1963)

The theory of a previous article, dealing with two types of two-level systems coupled to a loss mechanism (LM), is extended. The first extension consists of the consideration of the most general type of two-level system (TLS), in which the dipole moment is expanded in terms of the three Pauli spin matrices and unit matrix, the expansion coefficients being vectors (dipole vectors). The second extension consists of the addition to the thermal-reservoir type of LM of a large number of systems identical to the TLS under consideration. The TLS is described in terms of the time development of the Pauli matrices and differential equations are obtained for their expectation value in the presence of arbitrary driving fields. The Bloch equations for a magnetic dipole of spin $\frac{1}{2}$ are exhibited as a special case of these equations, corresponding to a particular combination of the dipole vectors. All other combinations describe electric dipole systems. Equations for two simple special cases of such systems are presented, one treated in the previous article and the other having permanent dipole moment. The frequency of oscillation of a freely decaying TLS is derived and shown to be shifted by an amount that depends on the relationship between the dipole vectors. It is pointed out that the commonly held belief that any TLS can be represented as a magnetic dipole of spin $\frac{1}{2}$ is only approximately correct in the presence of dissipation. The conditions under which the differential equations for the expectation values of the dynamical variables of the TLS can be converted into differential equations for macroscopic variables are discussed.

INTRODUCTION

IN a previous article,[1] it was pointed out that there are many important physical problems which may be considered, with some simplification, as that of a two-level system (TLS) coupled to a loss mechanism (LM), and a study was made of this problem. The LM was a thermal reservoir type of system, and two types of TLS were considered, the magnetic dipole type and the electric dipole type without permanent dipole moment that couples to its environment through one dynamical variable only. In the present article the analysis of I will be extended in two aspects. One extension is the consideration of the most general type of two-level dipole system, of which the two types considered in I are special cases.[2] The results will produce, on the one hand, a unified method of treating all two-level systems, and on the other hand, an explicit and systematic exhibition of differences among two-level systems. The other extension is the consideration of a more complicated LM, one which consists not only of a thermal reservoir but also—in addition to the reservoir—of a large number of systems identical to the TLS under consideration, loosely coupled to it, and surrounded by the same environment. (In the language of magnetic resonance, coupling of the "spin-spin" type as well as the "spin-lattice" type will be considered.) Part I and part II contain the first and second extensions, respectively. Some consequences of the results and several special cases are discussed in part III.

[1] I. R. Senitzky, Phys. Rev. **131**, 2827 (1963); hereafter referred to as I.

[2] The electric dipole TLS is far richer in possibilities than the magnetic dipole TLS, since the latter is restricted by the special properties of angular momentum. All cases other than the two considered in I are electric dipole types.

I.

The defining property of a dipole system may be given by the form of the interaction energy of the system with an electric or magnetic field,

$$H_{\text{int}}=-\mathbf{d}\cdot\mathbf{E},\ -\mathbf{d}\cdot\mathbf{H}, \tag{1}$$

respectively, where $\mathbf{d}$, the dipole moment, is an operator referring only to the system under consideration. Since our system is a TLS, the components of $\mathbf{d}$, that is d_x, d_y, and d_z are 2×2 Hermitian matrices, and for the most general TLS, these are arbitrary. Now, any 2×2 Hermitian matrix may be written as a linear superposition of the Pauli spin matrices together with the unit matrix. We may therefore write

$$\mathbf{d}=\mu\sum_{\alpha=1}^{4}\mathbf{a}_\alpha\sigma_\alpha, \tag{2}$$

where σ_4 is the unit matrix, σ_1, σ_2, and σ_3 are the three Pauli spin matrices, μ is a quantity having the dimension of dipole moment, and the $\mathbf{a}$'s are four real 3-dimensional vectors (determined by 12 real numbers, as many as determine three Hermitian 2×2 matrices) which determine completely the dipole moment operator of the TLS. The three Pauli spin matrices obey the well-known properties

$$\sigma_i^2=1,\quad \{\sigma_i,\sigma_j\}=0,\quad [\sigma_k,\sigma_l]=2i\sigma_m, \tag{3}$$

where $i\neq j$, $\{\ \}$ is the anticommutator bracket, and k, l, m stand for the cyclic permutation of 1, 2, 3. The $\mathbf{a}_i$'s may be normalized by the requirement that

$$d^2(\equiv\mathbf{d}\cdot\mathbf{d})=n\mu^2, \tag{4}$$

[the right side of Eq. (4) is to be considered multiplied

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

by the unit matrix] which gives

$$\sum_{\alpha=1}^{4} a_\alpha^2 = n, \tag{5}$$

where n is a number of the order of unity chosen for convenience in the particular case under consideration. The representation of the σ's is such that the energy of the TLS is given by $\frac{1}{2}\hbar\omega\sigma_3$. For an electric dipole system, nothing more can be said about the $\mathbf{a}$'s without additional information about the system. For a magnetic dipole TLS, however, the proportionality between magnetic moment and angular momentum requires that $\mathbf{a}_4=0$, and that $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ be perpendicular to each other and equal in magnitude.

The TLS is subject to fields produced by the LM and to externally imposed fields, $\mathbf{E}_{\rm LM}$ and $\mathbf{E}_{\rm ext}$, respectively, for the electric system, and $\mathbf{H}_{\rm LM}$ and $\mathbf{H}_{\rm ext}$ for magnetic system. We introduce the notation

$$\mathbf{F}=-(2/\hbar)\mu\mathbf{E}_{\rm LM}, \quad \mathbf{f}=-(2/\hbar)\mu\mathbf{E}_{\rm ext} \tag{6a}$$

for the electric dipole case, and

$$\mathbf{F}=-(2/\hbar)\mu\mathbf{H}_{\rm LM}, \quad \mathbf{f}=-(2/\hbar)\mu\mathbf{H}_{\rm ext} \tag{6b}$$

for the magnetic dipole case. The Hamiltonian may now be written, for either case, as

$$H=H_{\rm LM}+\tfrac{1}{2}\hbar\omega\sigma_3+\tfrac{1}{2}\hbar\textstyle\sum_\alpha \sigma_\alpha \mathbf{a}_\alpha\cdot\mathfrak{F}, \tag{7}$$

where

$$\mathfrak{F}=\mathbf{F}+\mathbf{f}. \tag{7a}$$

Comparison of Eq. (7) with Eqs. (I.60) and (I.70) shows that the electric dipole case in I corresponds to $a_\alpha=\delta_{\alpha 1}$ and the magnetic case corresponds to $\mathbf{a}_1=\hat{x}$, $\mathbf{a}_2=\hat{y}$, $\mathbf{a}_3=\hat{z}$, $\mathbf{a}_4=0$, $\hat{x}$, $\hat{y}$, $\hat{z}$ being unit Cartesian coordinate vectors.

A TLS can have only three linearly independent dynamical variables, since each variable corresponds to a 2×2 Hermitian matrix and, as mentioned previously, there are only three such linearly independent matrices possible, aside from the physically uninteresting unit matrix. In the present analysis, (as in I), the TLS will be described by the three Pauli spin matrices. Once we have these, we can obtain the dipole moment, the energy, and any other variable. The LM will be described (likewise as in I) by $H_{\rm LM}$ and $\mathbf{F}$.

The expression for the Hamiltonian, Eq. (7), may be rewritten as

$$H=H_{\rm LM}+\tfrac{1}{2}\hbar\omega\sigma_3+\tfrac{1}{2}\hbar\sum_{\alpha=1}^{4}\sum_{m=1}^{3}\sigma_\alpha a_{\alpha m}\mathfrak{F}_m, \tag{8}$$

where the vectors have been resolved along Cartesian axes labeled 1, 2, and 3. As in I, we consider F_1, F_2, and F_3 to act independently of one another; we assume that they refer to three independent but identical LM's. Correspondingly, we write

$$H_{\rm LM}=H_1+H_2+H_3. \tag{9}$$

(This separation of $H_{\rm LM}$ was not indicated explicitly in I, but is implicit in the argument used there.) The equations of motion for the TLS and LM, obtained from Eqs. (8) and (9), are

$$\dot{\sigma}_1=-\omega\sigma_2+\textstyle\sum_m(a_{2m}\mathfrak{F}_m\sigma_3-a_{3m}\mathfrak{F}_m\sigma_2), \tag{10a}$$

$$\dot{\sigma}_2=\omega\sigma_1+\textstyle\sum_m(a_{3m}\mathfrak{F}_m\sigma_1-a_{1m}\mathfrak{F}_m\sigma_3), \tag{10b}$$

$$\dot{\sigma}_3=\textstyle\sum_m(a_{1m}\mathfrak{F}_m\sigma_2-a_{2m}\mathfrak{F}_m\sigma_1), \tag{10c}$$

$$\dot{F}_m=-(i/\hbar)[F_m,H_m], \tag{10d}$$

$$\dot{H}_m=-(i/2)\textstyle\sum_\alpha a_{\alpha m}\sigma_\alpha[H_m,F_m]. \tag{10e}$$

The last two equations may be combined to give

$$\dot{F}_m(t)=-(i/\hbar)[F_m,H_m(0)]+(1/2\hbar)\sum_\alpha a_{\alpha m}\int_0^t dt_1 \times[F_m(t),[F_m(t_1),H_m(t_1)]\sigma_\alpha(t_1)], \tag{11}$$

which may be rewritten as an integral equation

$$F_m(t)=F_m^{(0)}(t)+(1/2\hbar)\sum_\alpha a_{\alpha m}\int_0^t dt_1\int_0^{t_1} dt_2\, U_m(t-t_1) \times[F_m(t_1),[F_m(t_2),H_m(t_2)]\sigma_\alpha(t_2)]U_m^{-1}(t-t_1), \tag{12}$$

where

$$U(\tau)=\exp[(i/\hbar)H_m(0)\tau], \tag{12a}$$

and where $F_m^{(0)}(t)$ is defined by

$$F_m^{(0)}(t)=F_m(0), \tag{12b}$$

$$\dot{F}_m^{(0)}(t)=-(i/\hbar)[F_m^{(0)}(t),H_m(0)]. \tag{12c}$$

The important properties of $F^{(0)}(t)$ needed for present purposes, and derived in I, are the following:

$$\langle F_m^{(0)}(t)\rangle=0, \tag{13}$$

$$\langle F_m^{(0)}(t_1)F_n^{(0)}(t_2)\rangle = \delta_{nm}(2/\pi)\int_0^\infty d\omega'[\eta(\omega')\cos\omega'(t_1-t_2) - i\xi(\omega')\sin\omega'(t_1-t_2)], \tag{14}$$

where

$$\xi(\omega')=\tfrac{1}{2}\pi\hbar Z^{-1}B(\omega')[1-\exp(-\hbar\omega'/kT)], \tag{14a}$$

$$\eta(\omega')=\tfrac{1}{2}\pi\hbar Z^{-1}B(\omega')[1+\exp(-\hbar\omega'/kT)], \tag{14b}$$

$$Z=\int_0^\infty dE\rho(E)\exp(-E/kT), \tag{14c}$$

$$B(\omega')=\int_0^\infty dE\rho(E+\hbar\omega')\rho(E) \times\tilde{F}^2(E+\hbar\omega',E)\exp(-E/kT), \tag{14d}$$

$\rho(E)$ being the density of energy states of the LM, $\tilde{F}^2(E_i,E_k)$ being the average over small ranges of E_i and E_k of $|F_{ik}^{(0)}|^2$, and T being the LM temperature. Both $\eta(\omega')$ and $\xi(\omega')$ are assumed to become vanishingly small

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

as ω' approaches zero, and approximately constant in the neighborhood of $\omega'=\omega$, the values in this neighborhood being denoted by η and ξ (without argument) respectively. The expression $\langle F^{(0)}(t_1)F^{(0)}(t_2)\rangle$ will occur, in the present analysis, as a factor in an integrand, the integration being over t_1, or t_2, or both. Consider the expression

$$\int_0^t dt_1\varphi(t_1)\langle F^{(0)}(t_1)F^{(0)}(t_2)\rangle \tag{15}$$

for $t\gg\omega^{-1}$. Substituting from Eq. (14) into (15), and interchanging the order of integration, one sees that if $\varphi(t_1)$ is approximately an oscillating function with angular frequency ω'', the main contribution to the integration comes from the neighborhood $\omega'\sim\omega''$. The same argument applies if the integration in (15) is over t_2 [with $\varphi(t_2)$ instead of $\varphi(t_1)$]. We may therefore write for use in subsequent integration

$$\langle F^{(0)}(t_1)F^{(0)}(t_2)\rangle \approx 2\left[\eta(\omega'')\delta(t_1-t_2)-\frac{i}{\pi}\xi(\omega'')\frac{\mathcal{P}}{t_1-t_2}\right], \tag{16a}$$

and

$$\langle\{F^{(0)}(t_1),F^{(0)}(t_2)\}\rangle\approx 4\eta(\omega'')\delta(t_1-t_2). \tag{16b}$$

The values of ω'' to be encountered in the present article will be only ω and 0. For the latter value, it is clear that $\langle F^{(0)}(t_1)F^{(0)}(t_2)\rangle=0$.

Equations (12) may be written as

$$F_m(t)=F_m{}^{(0)}(t)+\sum_\alpha a_{\alpha m}F_{\alpha m}{}^{(1)}(t), \tag{17}$$

where

$$F_{\alpha m}{}^{(1)}(t)=\frac{1}{2\hbar}\int_0^t dt_1\int_0^{t_1} dt_2 U_m(t-t_1) \times[F_m(t_1),[F_m(t_2),H_m(t_2)]\sigma_\alpha(t_2)] \times U_m{}^{-1}(t-t_1). \tag{17a}$$

By use of approximations based on the fact that the LM is affected only slightly by the TLS, it is shown in I that

$$F_{1m}{}^{(1)}(t)=-\bar\xi\sigma_2(t), \tag{18a}$$

$$F_{2m}{}^{(1)}(t)=\bar\xi\sigma_1(t), \tag{18b}$$

where $\bar\xi$ is a (c-number) function of t which is zero at $t=0$ and approaches the constant ξ in a time large compared to ω^{-1} but short compared to the time during which secular changes (in the TLS) take place. It is also shown in I [Eq. (I.76)] that

$$F_{3m}{}^{(1)}(t)=0, \tag{19a}$$

and by an identical argument[3] it can be shown that

$$F_{4m}{}^{(1)}(t)=0. \tag{19b}$$

[3] Equation (19a) follows from the fact that $\sigma_3(t)$ is a slowly varying function [while $\sigma_1(t)$ and $\sigma_2(t)$ are approximately oscillatory functions with (angular) frequency ω]. Equation (19b) follows from the fact that σ_4 is the unit matrix and does not vary at all.

Substituting from Eqs. (18) and (19) into (17), we obtain

$$F_m(t)=F_m{}^{(0)}(t)-a_{1m}\bar\xi\sigma_2(t)+a_{2m}\bar\xi\sigma_1(t). \tag{20}$$

This expression for $F_m(t)$ may now be substituted into Eqs. (10a)–(10c), the equations of motion for the TLS. Before we do that, however, it is convenient to rewrite the products of $\mathfrak{F}_m$ and σ_α in these equations as symmetrized products. (Note that $[F_m(t),\sigma_\alpha(t)]=0$.) Then, by substituting from Eq. (20) and utilizing the properties of the σ's given in Eq. (3), we get

$$\dot\sigma_1=-\omega\sigma_2+\sum_m(\tfrac{1}{2}a_{2m}\{\mathfrak{F}_m{}^{(0)},\sigma_3\} -\tfrac{1}{2}a_{3m}\{\mathfrak{F}_m{}^{(0)},\sigma_2\}+a_{1m}a_{3m}\bar\xi), \tag{21a}$$

$$\dot\sigma_2=\omega\sigma_1+\sum_m(\tfrac{1}{2}a_{3m}\{\mathfrak{F}_m{}^{(0)},\sigma_1\} -\tfrac{1}{2}a_{1m}\{\mathfrak{F}_m{}^{(0)},\sigma_3\}+a_{2m}a_{3m}\bar\xi), \tag{21b}$$

$$\dot\sigma_3=\sum_m[\tfrac{1}{2}a_{1m}\{\mathfrak{F}_m{}^{(0)},\sigma_2\}-\tfrac{1}{2}a_{2m}\{\mathfrak{F}_m{}^{(0)},\sigma_1\} -(a_{1m}{}^2+a_{2m}{}^2)\bar\xi], \tag{21c}$$

where

$$\mathfrak{F}_m{}^{(0)}\equiv F_m{}^{(0)}+f_m.$$

Equations (21) contain as unknowns, dynamical variables referring to the TLS only. $\mathbf{F}^{(0)}$ is a "prescribed" field, and so is, of course, $\mathbf{f}$. $\mathbf{F}^{(0)}$ is the quantum-mechanical version of a stochastic force. With $\bar\xi$ replaced by ξ, and with $f=0$, Eqs. (21) are of the type referred to in I as the Langevin equations for a TLS.

It should be noted that, while $F_m(t)$ and $\sigma_\alpha(t)$ commute, $F_m{}^{(0)}(t)$ and $\sigma_\alpha(t)$ do not. The commutation relations between the latter two operators may be obtained from Eq. (20), utilizing the fact that $[F_m(t),\sigma_\alpha(t)]=0$; the result is

$$[\sigma_1,F_m{}^{(0)}]=2ia_{1m}\bar\xi\sigma_3, \tag{22a}$$

$$[\sigma_2,F_m{}^{(0)}]=2ia_{2m}\bar\xi\sigma_3, \tag{22b}$$

$$[\sigma_3,F_m{}^{(0)}]=-2i\bar\xi(a_{1m}\sigma_1+a_{2m}\sigma_2). \tag{22c}$$

The problem may now be considered as formally defined by Eqs. (21)—Eqs. (22) being part of the prescription for $F^{(0)}$—with initial values for the σ's given by (their usual description when the Schrödinger picture is used)

$$\sigma_1(0)=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_2(0)=\begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \sigma_3(0)=\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{23}$$

From Eqs. (21) and (22) it can be shown, in a manner similar to that employed in I, that the solutions of Eqs. (22) subject to the initial conditions (23) are true spin operators, satisfying Eqs. (3). These solutions, of course, are operators both with respect to the TLS and LM. It is easy to see that this satisfactory state of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

affairs is due to the fact that $\bar{\xi}$ vanishes at $t=0$, since $F(0)$ and $\sigma_\alpha(0)$ are taken to be the uncoupled operators, and the coupling between TLS and LM is assumed to begin at $t=0$. Having shown the consistency of the formalism, we now replace $\bar{\xi}$ by ξ, as approximation for compulational purposes.

A solution of Eqs. (21) that is an operator in both TLS and LM spaces has not been found (even when all $\mathbf{a}$'s but $\mathbf{a}_1$ vanish, as in the electric dipole case in I). In order to make progress with Eqs. (21), we take expectation values in LM space. This does not simplify the situation automatically since expectation values of products, $\langle\{F_m{}^{(0)},\sigma_\alpha\}\rangle$, occur, and they are certainly not equal to the product of the expectation values. The evaluation of expectation values of these products will involve a significant approximation.

We consider $\mathfrak{F}^{(0)}$ (that is, both $F^{(0)}$ and f) to be a small quantity (compared to ω) of first order. Equations (21) can be rewritten as integral equations for the σ's in which the kernels are of first order. Defining the symbols $\Phi_1(t)$, $\Phi_2(t)$, $\Phi_3(t)$ by rewriting Eqs. (21) as

$$\sigma_1(t)=-\omega\sigma_2(t)+\Phi_1(t)\,, \tag{24a}$$

$$\dot{\sigma}_2(t)=\omega\sigma_1(t)+\Phi_2(t)\,, \tag{24b}$$

$$\dot{\sigma}_3(t)=\Phi_3(t)\,, \tag{24c}$$

we obtain, as equivalent integral equations,

$$\sigma_1=\sigma_1{}^{(0)}+\int_0^t dt_1\cos\omega(t-t_1)\Phi_1(t_1) -\int_0^t dt_1\sin\omega(t-t_1)\Phi_2(t_1)\,, \tag{25a}$$

$$\sigma_2=\sigma_2{}^{(0)}+\int_0^t dt_1\cos\omega(t-t_1)\Phi_2(t_1) +\int_0^t dt_1\sin\omega(t-t_1)\Phi_1(t_1)\,, \tag{25b}$$

$$\sigma_3=\sigma_3{}^{(0)}+\int_0^t dt_1\Phi_3(t_1)\,, \tag{25c}$$

where the $\sigma^{(0)}$'s are the solution for the free TLS satisfying the initial conditions. A substitution is now made from Eqs. (25) into all products $\{F_m{}^{(0)},\sigma_\alpha\}$ occurring in Eqs. (21). Thus

$$\begin{aligned}\langle\{F_m{}^{(0)}(t_1),\sigma_1(t_1)\}\rangle_{\rm LM}&\\ =\tfrac{1}{2}\sum_n\int_0^t dt_1\langle[a_{1n}\sin\omega(t-t_1)&\\ +a_{2n}\cos\omega(t-t_1)]\{F_m{}^{(0)}(t),\{\mathfrak{F}_n{}^{(0)}(t_1),\sigma_3(t_1)\}\}&\\ -a_{3n}\{F_m{}^{(0)}(t),\{\mathfrak{F}_n{}^{(0)}(t_1),\sigma_2(t_1)\}\}\cos\omega(t-t_1)&\\ -a_{3n}\{F_m{}^{(0)}(t),\{\mathfrak{F}_n{}^{(0)}(t_1),\sigma_1(t_1)\}\}&\\ \times\sin\omega(t-t_1)\rangle_{\rm LM}\,,&\end{aligned} \tag{26}$$

where use has been made of the fact that $\langle F_m{}^{(0)}\sigma_\alpha{}^{(0)}\rangle_{\rm LM}=\langle F_m{}^{(0)}\rangle\sigma_\alpha{}^{(0)}$, and of Eq. (13).[4] The significant approximation consists of ignoring the noncommutativity of σ_α and $F_m{}^{(0)}$ and replacing the product of the LM variables by its expectation value, both operations to be performed only in terms of higher order than the first. (See I for a discussion of this approximation.) Utilizing Eq. (13) and the Krönecker delta of Eq. (14), we obtain, with the above approximation,

$$\begin{aligned}\langle\{F_m{}^{(0)},\sigma_1\}\rangle_{\rm LM}&\\ =\int_0^t dt_1\langle\{F_m{}^{(0)}(t),F_m{}^{(0)}(t_1)\}\rangle&\\ \times\{\sigma_3(t_1)[a_{1m}\sin\omega(t-t_1)+a_{2m}\cos\omega(t-t_1)]&\\ -a_{3m}[\sigma_1(t_1)\sin\omega(t-t_1)+\sigma_2(t_1)\cos\omega(t-t_1)]\}\,.&\end{aligned} \tag{27}$$

The considerations related to Eqs. (15) and (16) may now be applied. $\sigma_3(t_1)$ is a slowly varying function of t_1, so that Eq. (16b) with $\omega''=\omega$ must be used for the integration associated with first square bracket in Eq. (27). The second square bracket is shown in I [Eq. (I.81); note that σ_1 and σ_2 approximately oscillate with frequency ω] to be itself a slowly varying function of t_1, so that Eq. (16b) with $\omega''=0$ must be used for the integration associated with the second square bracket, and the contribution of this integration vanishes. The result is

$$\langle\{F_m{}^{(0)},\sigma_1\}\rangle_{\rm LM}=2a_{2m}\eta\langle\sigma_3\rangle_{\rm LM}\,. \tag{28a}$$

In an entirely similar manner, one obtains

$$\langle\{F_m{}^{(0)},\sigma_2\}\rangle_{\rm LM}=-2a_{1m}\eta\langle\sigma_3\rangle_{\rm LM}\,, \tag{28b}$$

and

$$\langle\{F_m{}^{(0)},\sigma_3\}\rangle_{\rm LM}=2\eta(a_{1m}\langle\sigma_2\rangle_{\rm LM}-a_{2m}\langle\sigma_1\rangle_{\rm LM})\,. \tag{28c}$$

Substituting from Eqs. (28) into Eqs. (21), we have

$$\dot{\sigma}_1=(-\omega+\mathbf{a}_1\cdot\mathbf{a}_2\eta)\sigma_2-a_2{}^2\eta\sigma_1+\mathbf{a}_1\cdot\mathbf{a}_3\eta(\sigma_3-\sigma_0) +\mathbf{a}_2\cdot\mathbf{f}\sigma_3-\mathbf{a}_3\cdot\mathbf{f}\sigma_2\,, \tag{29a}$$

$$\dot{\sigma}_2=(\omega+\mathbf{a}_1\cdot\mathbf{a}_2\eta)\sigma_1-a_1{}^2\eta\sigma_2+\mathbf{a}_2\cdot\mathbf{a}_3\eta(\sigma_3-\sigma_0) +\mathbf{a}_3\cdot\mathbf{f}\sigma_1-\mathbf{a}_1\cdot\mathbf{f}\sigma_3\,, \tag{29b}$$

$$\dot{\sigma}_3=-(a_1{}^2+a_2{}^2)\eta(\sigma_3-\sigma_0)+\mathbf{a}_1\cdot\mathbf{f}\sigma_2-\mathbf{a}_2\cdot\mathbf{f}\sigma_1\,, \tag{29c}$$

where the notation

$$\sigma_0\equiv-\xi/\eta \tag{29d}$$

is used, and where LM expectation value brackets are omitted. (Since these are operator equations in TLS space, σ_0 is to be understood as being multiplied by the unit matrix.) It is easily seen that the results of I for the magnetic and electric dipole cases treated there are obtained by setting $a_{im}=\delta_{im}$ in the former case and $a_{im}=\delta_{i1}\delta_{m1}$ in the latter. Equations (29) describe the

[4] Where brackets are used to indicate expectation value with respect to the LM, the subscript "LM" will be added if the enclosed expression refers to both the TLS and LM, while no subscript will be used if the enclosed expression refers only to the LM.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

behavior of the most general TLS coupled to a thermal-reservoir type LM.

We can take expectation values with respect to the TLS in Eqs. (29). This leaves the equations formally unaltered, and the σ's now stand for their respective expectation values in both LM and TLS spaces. Henceforth, Eqs. (29) will be understood to refer to these expectation values. Discussion of the physical significance of these equations will be postponed until the second generalization has been carried out.

II.

The second generalization refers directly to the LM rather than to the TLS. Instead of assuming the LM to be only a thermal reservoir of given temperature, we consider the LM to consist of two parts: The first part (to be referred to as LM_1) is the same thermal reservoir considered heretofore; the second part (to be referred to as LM_2) consists of a large number of two-level systems identical to the TLS under consideration, situated in a similar environment, and loosely and randomly coupled to our system and to each other. The individual two-level systems of LM_2 are thus coupled (in addition to the coupling to each other and to the TLS under consideration) to a thermal reservoir and to external forces. The thermal reservoir may be LM_1 itself, but since the possibility of the transmission of effects through LM_1 will be excluded, it is simpler to consider the systems of LM_2 coupled to a thermal reservoir that is identical to LM_1 but independent. In the notation to be used, LM_2 consists of the mutually coupled two-level systems only, the thermal reservoir and external forces being described separately.

The Hamiltonian for our problem is obtained by adding several terms to the expression in Eq. (7), the result being

$$H=\tfrac{1}{2}\hbar\omega\sigma_3+\tfrac{1}{2}\hbar\sum_\alpha \sigma_\alpha \mathbf{a}_\alpha\cdot(\mathbf{F}^{(1)}+\mathbf{F}^{(2)}+\mathbf{f}) +H_{\mathrm{LM1}}+H_{\mathrm{LM2}}+H^{(2)}, \quad (30)$$

where H_{LM1} and H_{LM2} are the energy operators of LM_1 and LM_2, respectively, $F^{(1)}$ and $F^{(2)}$ are the coordinates through which LM_1 and LM_2, respectively, couple to the TLS, and $H^{(2)}$ contains the coupling term of LM_2 to any other systems, the energy operators for those systems, and the coupling term of LM_2 with prescribed external forces; that is, $H^{(2)}$ refers to everything that couples to LM_2 except the TLS under consideration.

We pass now from the Heisenberg picture, which has been used in the preceding discussion, to the interaction picture. All the operators are transformed into primed quantities in the manner[5]

$$A'=UAU^*, \quad (31)$$

where U is a unitary operator satisfying the differential equation and initial condition

$$i\hbar\dot{U}(t)=U(t)H^{(2)}(t), \quad (31a)$$

$$U(0)=1. \quad (31b)$$

All the primed operators now satisfy the equation of motion

$$i\hbar\dot{A}'=[A', H'-H^{(2)\prime}]+i\hbar(\partial A'/\partial t); \quad (32)$$

in other words, the effective Hamiltonian that enters into the equations of motion for the primed operators no longer contains any reference to the influence of external systems or forces on LM_2. Formally, LM_1 and LM_2 appear equivalently in these equations. In this transformation, the effect of the external influences on LM_2 has been removed from the operators and transferred to the state vector, which becomes

$$\Psi'(t)=U(t)\Psi. \quad (33)$$

It should be remembered that Ψ is independent of the time and would be the initial state if the Schrödinger picture were used; Eq. (33) may be written as

$$\Psi'(t)=U(t)\Psi'(0). \quad (34)$$

Now, if there were no coupling between the TLS and LM_2, Eq. (34) would yield the entire time development of LM_2, and leave all the other systems to which reference is made in the Hamiltonian of Eq. (30) unaffected, that is in their Heisenberg (initial) state. The presence of the $\frac{1}{2}\hbar\sigma_\alpha\mathbf{a}_\alpha\cdot\mathbf{F}^{(2)}$ coupling term affects U only indirectly through the time development of $H^{(2)}(t)$ [note that $H^{(2)}(0)$ does not contain this coupling term] and does not produce lowest order effects. The difference between $\Psi'(t)$ and the corresponding state in which there is no coupling between the TLS and LM_2 is due only to the effect on LM_2 of the TLS, and is slight. We approximate by ignoring this effect in $\Psi'(t)$ and consider $\Psi'(t)$ to denote the state in which the TLS and LM_1 are unaffected while LM_2 has developed under the influence of those systems (and external forces) explicitly entering into $H^{(2)}$. (This approximation is similar in spirit to approximations previously made for LM_1, where LM variables in interaction terms were replaced by their uncoupled values. In the present instance, LM_2 may be uncoupled only from the TLS but not from the effect of $H^{(2)}$. This is the essential difference between LM_1 and LM_2.)

Analytically, the approximation just performed may be described as follows: Set

$$\Psi=\psi_{\mathrm{TLS}}\psi_{\mathrm{LM1}}\psi_{\mathrm{LM2}}, \quad (35a)$$

(this implies that the coupling is turned on at $t=0$), which is the same as

$$\Psi'(0)=\psi_{\mathrm{TLS}}'(0)\psi_{\mathrm{LM1}}'(0)\psi_{\mathrm{LM2}}'(0). \quad (35b)$$

Consider $H^{(2)}(t)$ to be obtained from $H^{(2)}(0)$ (by means of the Heisenberg equations of motion) by neglecting

[5] It should be noted that the interaction picture is obtained here by starting with the Heisenberg picture. This accounts for the order of the factors in the right side of Eq. (31a), the opposite of that used when starting with the Schrödinger picture.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the $F^{(2)}$ term in Eq. (30). Then $H^{(2)}(t)$ and, therefore—as is evident from Eq. (31a)—$U(t)$ contain only LM_2 variables. Thus,

$$\Psi'(t)=\psi_{TLS}'(0)\psi_{LM1}'(0)U(t)\psi_{LM2}'(0)$$
$$=\psi_{TLS}'(0)\psi_{LM1}'(0)\psi_{LM2}'(t). \quad (36)$$

Since LM_1 and LM_2 are large and incompletely described systems, we will describe their initial states, as has been done previously for LM_1, by ensemble averages. For this purpose, it is more convenient to use density matrices instead of state vectors. We therefore write, in place of Eq. (36),

$$P'(t)=\rho_{TLS}'(0)\rho_{LM1}'(0)\rho_{LM2}'(t), \quad (37)$$

where P is the density matrix for the combination of systems, and the ρ's are individual density matrices. Both $\psi_{LM2}'(t)$ and $\rho_{LM2}'(t)$ describe the development of LM_2 in the absence of coupling to the TLS under consideration.

We come now to the essential aspect, and significant approximation, of the analysis of LM_2. LM_2 is obviously not a thermal reservoir, since the individual two-level systems of which it is composed can interchange an amount of energy comparable to their own energy with the thermal reservoir and external forces to which they are coupled. LM_2 cannot, therefore, be described by a fixed temperature. We assume, however, that the randomness associated with LM_2 (due, mainly, to the random coupling among the individual two-level systems) is sufficiently great so that at any given time LM_2 may be described by a temperature, this temperature being, in general, a function of the time.[6] According to this assumption, the density matrix for LM_2 is given by

$$\rho_{LM2}'(t)=Z_2^{-1}\exp[-H_{LM2}/kT(t)], \quad (38)$$

where

$$Z_2(t)=\text{Trace}\exp[-H_{LM2}/kT(t)]. \quad (38a)$$

$T(t)$ is left unspecified for the time being, but we assume that it changes slowly compared to $\exp(i\omega t)$.

Now, LM_2, like LM_1, is a large system with many energy levels closely spaced. It may therefore be treated, in the analysis of its interaction with the TLS, in the same manner as LM_1, but with cognizance of the time dependence of the temperature. (It is to be noted that the temperature enters into the analysis only through the density matrix.) Our problem thus reduces itself to the analysis of the interaction of a TLS with external fields and with two LM's, one of which has a time-dependent temperature.

Before this problem is considered, it is useful to look at a simpler situation, one in which there are two LM's with different, but *fixed*, temperatures. Very little analysis is required to notice that the change required in the results obtained for a single LM is the replacement of η by $\eta^{(1)}+\eta^{(2)}$, and the replacement of ξ by $\xi^{(1)}+\xi^{(2)}$, where the superscripts (1) and (2) refer to LM_1 and LM_2, respectively, and $\xi^{(i)}$ and $\eta^{(i)}$ are defined for each LM in the same manner as ξ and η, respectively, by Eqs. (14a) and (14b). From Eq. (29d) we have

$$\sigma_0^{(i)}=-\xi^{(i)}/\eta^{(i)}; \quad (39)$$

thus Eqs. (29a)–(29c) remain valid provided we set

$$\eta=\eta^{(1)}+\eta^{(2)}, \quad (40a)$$

and

$$\sigma_0=(\sigma_0^{(1)}\eta^{(1)}+\sigma_0^{(2)}\eta^{(2)})(\eta^{(1)}+\eta^{(2)})^{-1} \quad (40b)$$

We return now to situation in which $T^{(2)}$ is time-dependent. This time dependence affects several calculations that lead to Eqs. (29). First, there is the computation of expectation values of LM_2 variables. The expectation value of an operator $A'(t)$ referring to a single time (and uncoupled from the TLS) poses no problem since it is given by Trace $\rho_{LM}'(t)A'(t)$. The expectation value of an operator of the type $\{A_1'(t_1),A_2'(t_2)\}$, referring to two different times, requires, in general, another method of evaluation. In the present analysis, however, this type of expectation value is needed only in cases where the result is negligible unless t_1 is very close to t_2. [See, for instance, Eqs. (14) and (16).] Bearing in mind that $T^{(2)}(t)$ is a slowly varying function, we can use previous results for this type of expectation value, with $\xi^{(2)}$ and $\eta^{(2)}$—which are functions of $T^{(2)}$—evaluated at either t_1, t_2, or an intermediate value. Then, there is the question of time integration where $\xi^{(2)}$ or $\eta^{(2)}$ is part of the integrand. [See, for instance, Eq. (27).] Here too, no formal modification of the final result is necessary, for in this integration the main contribution comes only from a very small neighborhood of a definite time, so that $\xi^{(2)}$ or $\eta^{(2)}$ may be evaluated at that time and taken outside the integral. It is seen, therefore, that Eqs. (29a)–(29c) and (40) remain formally unchanged, but with $\eta^{(2)}$ and $\sigma_0^{(2)}$ being considered now functions of t (the same argument as that of the σ's and **f**).

Our final task is the evaluation of $T^{(2)}$, or, more directly, of $\sigma_0^{(2)}$ and $\eta^{(2)}$, through which the behavior of the TLS depends on $T^{(2)}$. From Eqs. (39), (14a), and (14b), we have

$$\sigma_0^{(2)}=\frac{\exp(-\hbar\omega/kT^{(2)})-1}{\exp(-\hbar\omega/kT^{(2)})+1}. \quad (41)$$

This is the expectation value of the energy, in units of $\frac{1}{2}\hbar\omega$, of a TLS of the type under consideration when it is in thermal equilibrium with a reservoir at temperature $T^{(2)}$. We may therefore regard $\sigma_0^{(2)}$ as the average energy (at time t) of the two-level systems that constitute LM_2. Since these two-level systems are identical to the TLS under consideration and are situated in a

[6] Since LM_2 is composed of two-level systems, its temperature may assume negative as well as positive values, depending on the average energy of the two-level systems.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

similar environment, this energy is also equal to $\sigma_3(t)$.[7] Thus, we have the important result

$$\sigma_0^{(2)}=\sigma_3(t). \tag{42}$$

As far as $\eta^{(2)}$ is concerned, no such simple consideration applies. The best that can be done within the scope of the present analysis is to observe that $\eta^{(2)}$ is not as strongly dependent on $T^{(2)}$ as $\sigma_0^{(2)}$ (or $\xi^{(2)}$)—as can be seen from Eqs. (14)—and to approximate it by an average value.

From Eqs. (29a)–(29c), (40), and (42) we finally obtain

$$\dot\sigma_1=(-\omega+\tilde\eta\mathbf{a}_1\cdot\mathbf{a}_2)\sigma_2-\tilde\eta a_2^2\sigma_1+\eta\mathbf{a}_1\cdot\mathbf{a}_3(\sigma_3-\sigma_0)+\mathbf{a}_2\cdot\mathbf{f}\sigma_3-\mathbf{a}_3\cdot\mathbf{f}\sigma_2, \tag{43a}$$

$$\dot\sigma_2=(\omega+\tilde\eta\mathbf{a}_1\cdot\mathbf{a}_2)\sigma_1-\tilde\eta a_1^2\sigma_2+\eta\mathbf{a}_2\cdot\mathbf{a}_3(\sigma_3-\sigma_0)+\mathbf{a}_3\cdot\mathbf{f}\sigma_1-\mathbf{a}_1\cdot\mathbf{f}\sigma_3, \tag{43b}$$

$$\dot\sigma_3=-(a_1^2+a_2^2)\eta(\sigma_3-\sigma_0)+\mathbf{a}_1\cdot\mathbf{f}\sigma_2-\mathbf{a}_2\cdot\mathbf{f}\sigma_1, \tag{43c}$$

where

$$\tilde\eta\equiv\eta^{(1)}+\eta^{(2)},\quad \eta\equiv\eta^{(1)},\quad \sigma_0\equiv\sigma_0^{(1)}. \tag{43d}$$

In summary, these equations describe the behavior (in terms of expectation values) of a general TLS subject to external fields and coupled both to a thermal reservoir and a large number of systems identical to itself. The dipole moment is given in terms of the σ's by Eq. (2), and the energy is $\frac{1}{2}\hbar\omega\sigma_3$. Two relaxation constants, $\tilde\eta$ and η (with η being a better "constant" than $\tilde\eta$, which may be regarded as depending to some extent on the energy of the TLS) and an equilibrium energy σ_0 enter into these equations. The equilibrium energy is that associated with the temperature T_1 of the thermal reservoir, and is obtained by replacing superscript (2) with superscript (1) in Eq. (41).[8]

III.

A few special types of TLS are of immediate interest. As mentioned previously, the magnetic dipole type has $\mathbf{a}_1=\hat x$, $\mathbf{a}_2=\hat y$, $\mathbf{a}_3=\hat z$, $\mathbf{a}_4=0$. Equations (43) become

$$\dot\sigma_x=-\omega\sigma_y-\tilde\eta\sigma_x+f_x\sigma_z-f_z\sigma_x, \tag{44a}$$

$$\dot\sigma_y=\omega\sigma_x-\tilde\eta\sigma_y+f_z\sigma_x-f_x\sigma_z, \tag{44b}$$

$$\dot\sigma_z=-2\eta(\sigma_3-\sigma_0)+f_x\sigma_y-f_y\sigma_x. \tag{44c}$$

These are essentially the Bloch equations,[9] with $(2\eta)^{-1}$ being the longitudinal relaxation time T_1, and $\tilde\eta^{-1}$ being the transverse relaxation time T_2. If the spin-lattice coupling is regarded as coupling between the TLS and a thermal reservoir, and spin-spin coupling as coupling among identical two-level systems, then Eqs. (44) and (43d) show the contribution of each type of coupling to the relaxation process. It should be noticed that it is not, in general, correct to refer to T_2 as the spin-spin relaxation time (as is often done) unless $\eta^{(1)}$ is negligible compared to $\eta^{(2)}$. If $\eta^{(2)}$ is zero, Eqs. (44) reduce to the results of I [Eqs. (I.84)].

As far as the electric dipole TLS is concerned, there are many possible types, depending on the choice of the $\mathbf{a}$'s. $\mathbf{a}_4$ does not enter into the equations of motion and, therefore, does not affect the behavior of the TLS. $\mathbf{a}_4$ and $\mathbf{a}_3$ determine the "permanent" dipole moment, or the nonoscillating part of the (expectation value of the) dipole moment of the free TLS, for only σ_3 can have a nonoscillating value different from zero when $f=\tilde\eta=\eta=0$ in Eqs. (43). Since $\mathbf{a}_4$ plays no dynamic role, we will ignore it henceforth, and refer to the constant part of σ_3 (times $\mu\mathbf{a}_3$) as the permanent dipole moment. A case in which the dipole moment has a particularly simple appearance is that in which $\mathbf{a}_2=\mathbf{a}_3=0$, so that $\mathbf{d}=\mu\mathbf{a}_1\sigma_1$. This is the electric dipole case treated in I. Taking $a_1^2=1$, $\mathbf{a}_1\cdot\mathbf{f}\equiv f$, Eqs. (43) become

$$\dot\sigma_1=-\omega\sigma_2, \tag{45a}$$

$$\dot\sigma_2=\omega\sigma_1-\tilde\eta\sigma_2-f\sigma_3, \tag{45b}$$

$$\dot\sigma_3=-\eta(\sigma_3-\sigma_0)+f\sigma_2, \tag{45c}$$

which, for $\eta^{(2)}=0$, reduce to Eqs. (I.67). A permanent dipole moment may be added to this case, while still maintaining a relatively simple form of the equations, by considering $\mathbf{a}_2=0$, $\mathbf{a}_3=\epsilon\mathbf{a}_1$. (Here, the permanent dipole moment has the same spatial direction as the oscillating dipole moment.) The result is

$$\dot\sigma_1=-\omega\sigma_2-\epsilon\dot\sigma_3, \tag{46a}$$

$$\dot\sigma_2=\omega\sigma_1-\tilde\eta\sigma_2+\epsilon f\sigma_1-f\sigma_3, \tag{46b}$$

$$\dot\sigma_3=-\eta(\sigma_3-\sigma_0)+f\sigma_2, \tag{46c}$$

where Eq. (46a) has been simplified in appearance by substitution from Eq. (46c).

It is not the purpose of the present article to present equations for large number of special cases, although there are many others than the above three particularly simple ones that are of interest. Neither is it the purpose of the present article to discuss the solution of the above equations for given driving fields, which will be done in a forthcoming article. We will, however, discuss the solution of the equations in the absence of a driving field, that is, the free decay of a TLS from given initial conditions. For the sake of simplicity, we consider only

[7] Strictly speaking, there is an approximation involved in this statement, since LM expectation values are averages both with respect to a thermal (canonical) ensemble—implicit in the density matrix employed—and a quantum-mechanical ensemble, while TLS expectation values are averages only with respect to the latter.

[8] Equations (43), involving two dissipation parameters, $\tilde\eta$ and η, are more general than the LM considered in the above discussion. More general loss mechanisms may be obtained by replacing LM_1 with several thermal reservoirs at different temperatures, some of these temperatures being, possibly, prescribed slowly varying functions of the time. Equations of the form of Eq. (40) can then be used to obtain new values for the two dissipation parameters with Eqs. (43) remaining formally unchanged.

[9] F. Bloch, Phys. Rev. **70**, 460 (1946); R. K. Wangsness and F. Bloch, *ibid.* **89**, 728 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the case in which $\mathbf{a}_3$ is either perpendicular to both $\mathbf{a}_1$ and $\mathbf{a}_2$, or is equal to zero. From Eqs. (43) we have, under the above conditions,

$$\sigma_1=(-\omega+\tilde{\eta}\mathbf{a}_1\cdot\mathbf{a}_2)\sigma_2-\tilde{\eta}a_2^2\sigma_1, \tag{47a}$$

$$\dot{\sigma}_2=(\omega+\tilde{\eta}\mathbf{a}_1\cdot\mathbf{a}_2)\sigma_1-\tilde{\eta}a_1^2\sigma_2, \tag{47b}$$

$$\dot{\sigma}_3=-(a_1^2+a_2^2)\eta(\sigma_3-\sigma_0). \tag{47c}$$

The last equation may be solved immediately, the result being

$$\sigma_3=\sigma_0+[\sigma_3(0)-\sigma_0]\exp[-(a_1^2+a_2^2)\eta t]. \tag{48}$$

Equations (47a) and (47b) yield the same equation for both σ_1 and σ_2:

$$\ddot{\sigma}_{1,2}+\tilde{\eta}(a_1^2+a_2^2)\dot{\sigma}_{1,2}+\{\omega^2+\tilde{\eta}^2[a_1^2a_2^2-(\mathbf{a}_1\cdot\mathbf{a}_2)^2]\}\sigma_{1,2}=0. \tag{49}$$

The solution of Eq. (49) is

$$\sigma_1=\exp[-\tilde{\eta}(a_1^2+a_2^2)t][a_1\cos\Omega t+B_1\sin\Omega t], \tag{50a}$$

$$\sigma_2=\exp[-\tilde{\eta}(a_1^2+a_2^2)t][a_2\cos\Omega t+B_2\sin\Omega t], \tag{50b}$$

where A_i and B_i are constants determined by the initial conditions, and

$$\Omega^2=\omega^2\{1-(\tilde{\eta}^2/\omega^2)[(\mathbf{a}_1\cdot\mathbf{a}_2)^2+\tfrac{1}{4}(a_1^2-a_2^2)^2]\}. \tag{50c}$$

We see that σ_3 approaches exponentially its equilibrium value, and that σ_1 and σ_2 undergo exponentially damped oscillation. The frequency of oscillation Ω is of interest. For $\mathbf{a}_1$ and $\mathbf{a}_2$ equal in magnitude and perpendicular to each other, Ω is equal to ω, the frequency of the free TLS; otherwise Ω is less than ω. Equation (50c) puts into perspective the results of I, where it was found that the frequency of the undriven damped oscillator was equal to that of the free oscillator for the magnetic dipole case but not for the electric dipole case. It is seen, now, that the magnetic dipole TLS has just that combination of $\mathbf{a}$'s which yields $\Omega=\omega$, but the electric dipole TLS treated in I has $\mathbf{a}_2=0$, $a_1^2=1$, giving

$$\Omega^2=\omega^2[1-\tfrac{1}{4}(\tilde{\eta}^2/\omega^2)], \tag{51}$$

in accordance with the results of I.[10]

One encounters occasionally the statement that any TLS is equivalent to a system of spin $\frac{1}{2}$ in an appropriate magnetic field.[11] While this statement is true in the absence of dissipation, it is only an approximation when the TLS is coupled to an LM, as the above frequency consideration reveals, and as can be seen from the dependence of Eqs. (43) on the type of TLS considered.

It was shown previously [Eqs. (44)] that the Bloch equations are a special case of Eqs. (43). The Bloch equations were developed originally for macroscopic matter. They also apply to the expectation values for a microscopic system, and, in this sense, correspond to a special case of Eqs. (43). One might ask how the present theory applies to macroscopic matter, or to a large number of systems identical to—but possibly, with different orientation than—the TLS under consideration. Can one, as in the case of the Bloch equations, regard a linear combination of the (four) σ's as the component of dipole moment along a given direction for macroscopic matter containing N systems similar to our TLS? The macroscopic dipole moment $\mathbf{D}$ is given by

$$\mathbf{D}=\sum_{j=1}^{N}\mathbf{d}^{(j)}=\sum_{j=1}^{N}\sum_{\alpha=1}^{4}\mathbf{a}_\alpha^{(j)}\sigma_\alpha^{(j)}, \tag{52}$$

where the superscript j refers to the jth TLS. The above question is equivalent to asking if $\sigma_\alpha^{(j)}$ is independent (except for a multiplicative constant) of j. It is clear that if the systems are all oriented in the same manner, then $\mathbf{a}_\alpha^{(j)}=\mathbf{a}_\alpha$, $\sigma_\alpha^{(j)}=\sigma_\alpha$,

$$\mathbf{D}=N\mathbf{d}, \tag{53}$$

and the answer to the above question is, obviously, in the affirmative. If the microscopic systems are not all oriented similarly, then the coefficients $\mathbf{a}_i^{(j)}\cdot\mathbf{f}$ in Eq. (43a) are different for differently oriented systems. If, because of symmetry conditions, the solutions of Eqs. (43) are essentially unaffected by a particular distribution of orientations, then the answer is still in the affirmative, but if no such symmetry exists, then Eqs. (43) cannot be regarded as equations for macroscopic polarization. The macroscopic result must be obtained through Eq. (52), that is, by solving first the microscopic problem, and then superposing the solutions.

The essential aspects of the orientation of a magnetic dipole TLS are determined by the field that is responsible for the level separation. If this field is the same for all the two-level systems, Eqs. (43)—which reduce in this instance to the Bloch equations—may be regarded as equations for macroscopic polarization. Such is the case for macroscopic matter in an external (laboratory) field. If, however, the field responsible for level separation is an internal field and has different orientation for different dipoles, the above equations may be applied only on a microscopic scale.

ACKNOWLEDGMENT

The author wishes to thank Professor Julian Schwinger for illuminating discussion and for his reading of the manuscript.

[10] A comparison might be made between Eqs. (47a) and (47b) on the one hand, and equations for the coordinate and momentum of a harmonic oscillator with dissipation on the other. If both coordinate and momentum couple to the loss mechanism, then the equations of motion for the harmonic oscillator variables (in appropriate units) are

$$\dot{p}=-\omega q-\eta_1 p, \quad \dot{q}=\omega p-\eta_2 q.$$

It is seen that if $\mathbf{a}_1$ and $\mathbf{a}_2$ are orthogonal, σ_2 and σ_1 correspond to the coordinate and momentum, respectively, of the harmonic oscillator. It can also be seen readily that the oscillatory frequency of the freely decaying harmonic oscillator is equal to ω only if $\eta_1=\eta_2$. Comparison with the dissipation terms in Eqs. (47a) and (47b) shows that this corresponds to our requirement (in addition to that of orthogonality) that $a_1^2=a_2^2$.

[11] N. Bloembergen and Y. R. Shen, Phys. Rev. 133, A37 (1964); A. Abragam, *The Principles of Nuclear Magnetism* (Oxford University Press, New York, 1961), p. 36.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cluster Expansion in the Heitler–London Approach to Many-Electron Problems*

TADASHI ARAI
Argonne National Laboratory, Argonne, Illinois
(Received 2 January 1964)

The Heitler–London method based on nonorthogonal atomic orbitals is applied to arrays of an infinitely large number of atoms. The electronic energy is given by a quotient with strongly divergent numerator and denominator, which until now has defied correct computation. By noticing some resemblance of this problem to the linked-cluster expansion in many-body problems, we have now developed a new method to compute the Heitler–London energy. Here, the numerator and denominator are divided simultaneously by a common factor, which leads to a set of recurrence relations between the normalization matrices **F**. The matrices **F** are essential parts of the quotient to be calculated. When an overlap integral $\langle k|h\rangle$ is represented by a line starting from h and ending at k, the calculation of **F** using the recurrence relations is carried out systematically by drawing diagrams consisting of connected loops. Since our present aim in applying the Heitler–London method is to compute spin-wave spectra, the calculation is carried out in the complete space of spin waves and the energy expression is given by a Hermitian matrix. This introduces additional matrices **A**. The computation of **A** is also carried out by the diagram technique, since **A** can be expanded into an asymptotic series using **F**. Finally the energy matrix is written as a sum over connected diagrams, in accordance with the speculation obtained from the linked-cluster expansion. An error introduced by truncating the series of diagrams is also calculated. The present method not only ensures that the energy density in the Heitler–London method is finite, but also provides an accurate and practical way to compute the Heitler–London energy, which has never been accomplished previously. To calculate the ground-state energy the technique is simplified since the computation of **A** can be completely eliminated.

I. INTRODUCTION

FOR many years, the concept of exchange coupling has played an important role in theories of chemical bond and magnetism. This concept was first introduced in the Heitler–London theory[1] of molecular binding and applied by Heisenberg[2] to ferromagnetism. Although the Heisenberg theory has been formulated in a very attractive form by Van Vleck[3] and Møller[4] and much theoretical work in magnetism has used this formalism, the method has been criticized in many respects.

In this paper, we will limit our discussion to one of the difficulties inherited in the Heitler–London approach to magnetic problems; that is, the nonorthogonality difficulty in calculating energy spectra. The other aspects of the difficulties will be found in, for example, review articles by Herring.[5]

The nonorthogonality difficulty can be stated as follows. If we assume that there are no spin-orbit terms in the Hamiltonian, the calculation of the Heitler–London energy of the form

$$E=\int\Psi^*\mathcal{H}\Psi d\tau\Big/\int\Psi^*\Psi d\tau \tag{1.1}$$

is reduced to that of the Heisenberg exchange Hamiltonian

$$\mathcal{H}_{\rm eff}=-2\sum_{i>j}J_{ij}\mathbf{S}_i\cdot\mathbf{S}_j, \tag{1.2}$$

provided the wave function Ψ of this system is constructed from *orthogonal* atomic orbitals.

The orthogonality of atomic orbitals makes the exchange integrals J_{ij} thus obtained all positive. Since a positive exchange integral in the Heisenberg theory corresponds to ferromagnetism, this approach leads to the obviously absurd conclusion that all substances are ferromagnetic.

To utilize the Heitler–London method to full advantage, nonorthogonal atomic orbitals should be used as the basis. However, the calculation of energy spectra becomes extremely difficult since both numerator and denominator of the energy expression (1.1) diverge when the number N of electrons involved increases. It is not at all obvious if it is permissible to discard, from the energy expression, terms of higher order in overlaps. Such terms may appear N times, N^2 times, or even more. Slater[6] and Inglis[7] pointed out this difficulty, which led to grave doubts as to the validity of the Heisenberg exchange Hamiltonian[8,9] (1.2).

Previously,[10] we have discussed this difficulty. (This paper will be referred to as paper I.) The remaining problems will also be examined in detail and will be published elsewhere.[11] The conclusions will be summarized as follows: The exact calculation of the Heitler–London energy can be converted to the eigenvalue problem of a spin Hamiltonian of the form

$$\mathcal{H}_{\rm spin}=\sum_P(-1)^P J_P P^\sigma, \tag{1.3}$$

where P is a permutation operator and P^σ the corresponding permutation of spin variables. The exchange energy J_P for a single interchange P of electrons is very nearly equal to the value one would compute by neglecting terms of higher order in overlaps. The con-

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

[1] W. Heitler and F. London, Z. Physik 44, 455 (1927).
[2] W. Heisenberg, Z. Physik 49, 619 (1928).
[3] J. H. Van Vleck, *Theory of Electric and Magnetic Susceptibilities* (Clarendon Press, Oxford, 1932).
[4] C. Møller, Z. Physik 82, 559 (1933).
[5] C. Herring, Rev. Mod. Phys. 34, 631 (1962); also J. Appl. Phys. 31, 3S (1960).
[6] J. C. Slater, Phys. Rev. 35, 509 (1930).
[7] D. R. Inglis, Phys. Rev. 46, 135 (1934).
[8] J. C. Slater, Rev. Mod. Phys. 25, 199 (1953).
[9] G. Heber, Fortschr. Physik 1, 707 (1954).
[10] T. Arai, Phys. Rev. 126, 471 (1962).
[11] T. Arai and C. Herring (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tribution from the higher order terms in Hamiltonian $\sum_P'(-1)^P J_P P^\sigma$, where the summation $\sum_P'$ excludes all single interchanges, is small.

This gives a rigorous justification for using the Heisenberg exchange Hamiltonian in the calculation of thermodynamical functions.[12] However, we have not succeeded in evaluating the value of J_P when P is not a single interchange. Although we have bounded the total energy E as well as the exchange energy J_P for a single interchange by applying algebraic identities that exist between J_P's, the values of the bounds are merely first order in the approximation. The method cannot be extended immediately to narrow the bounds beyond the single interchange approximation. This implies that the accuracy of the Heitler–London method so far is limited up to that of the Heisenberg model.

If we would like to use the Heitler–London method as a rigorous mathematical tool to calculate many-electron systems, we have to provide an expansion technique with which the energy expression (1.1) as well as the expectation value of any observable quantity can be calculated up to a desired accuracy. This is the purpose of the present paper.

To achieve this goal, we have to carry out more detailed calculations of both the numerator and the denominator of (1.1). It is, of course, impossible to proceed with the calculation directly. However, we find that our problem has some resemblance to the linked cluster expansion for Green functions.[13] In either case, the quantity to be calculated is a quotient with divergent denominator and numerator. Use of Dyson's equation[14] reduces the numerator of a Green function to a factor multiplying all connected diagrams. Since the factor is exactly equal to the denominator and cancelled by it, the Green function consists of just the sum over all connected diagrams.

In the Heitler–London calculation, a set of recurrence relations (3.14) and (3.15) for the **F** matrices will take the place of Dyson's equation. Here the **F** matrices defined by (3.1) are the normalization matrices and the essential quantities to be calculated. Since the recurrence relations reduces the **F** matrices for N-electron systems to those for $N-1, N-2, \cdots$ electron systems, the successive application of the relations leads to the complete expansion of the **F** matrices in terms of overlaps. The expansion technique is similar to that used in both classical and quantum-theoretical treatments of many-body problems, and use of diagrams will facilitate the computation. It will be finally found that **F** is written as a sum over connected diagrams, when an overlap integral of the type $\langle\chi_k|\chi_h\rangle$ is denoted by a line starting from the center h of the atomic orbital χ_h and ending at k of χ_k. Following the prescriptions (i)–(iv) described in Sec. III(1b), diagrams for **F** can be drawn without knowledge of the expansion technique used. An error accompanied by truncating the expansion will be calculated by using a theorem which will be probed in the Appendix.

Since the energy matrix should be Hermitian, it is necessary to calculate, in addition to **F**, the matrix **A** defined by (3.2). The matrix **A** is to correct the deviation of **F** from Hermitian form. Although **A** is close to a unit matrix, the straight forward calculation is very difficult. In Sec. III(2), we shall, instead, develop an asymptotic expansion of **A** and discuss the convergence of the series. The calculation can also be carried out automatically by drawing diagrams.

By multiplying **F** by **A**, the energy matrix can be obtained and it will be found that the energy matrix is also written as a sum over connected diagrams.

The significance of the present method will be summarized as follows.

(1) By using this method, a rigorous calculation of the Heitler–London method including all overlaps can be carried out without difficulty. The effect of overlaps between distant atoms has been cancelled out rigorously. Such a calculation has never been performed for a many-electron system previously.[15,16] In fact, this has been considered impossible from the mathematical point of view.[8]

(2) The awesome task of calculating the inverse of the overlap matrix[17] is completely eliminated. The calculation can be carried out just by drawing diagrams. Even for a system with a small number of electrons, therefore, this method will be more convenient than inverting the overlap matrix directly.

(3) The fact that the energy matrix consists of just a sum over connected diagrams implies that the energy density is a finite number. This will, in practice, eliminate the nonorthogonality difficulties. Actually, a rigorous proof for this problem will be given in a forthcoming paper.[11]

So far, the calculation was carried out in the total space of the Heitler–London wave functions aiming at the computation of spin-wave spectra. The wave functions are naturally described by a vector in that space and the energy matrix is given in the form (2.26) or (2.27) rather than (1.1). This is why the matrix **A** appears. If, on the other hand, one is interested in the

[12] For a general review of the theory of direct exchange see C. Herring, in *Magnetism*, edited by G. Rado and H. Suhl (Academic Press Inc., New York, to be published).

[13] For instance, P. Nozières and D. Pines, Nuovo Cimento **9**, 47 (1958). J. Hubbard, Proc. Roy. Soc. (London) **A240**, 539 (1957) and **A243**, 336 (1957).

[14] F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[15] The only exception is the calculation by F. Takano, J. Phys. Soc. Japan **14**, 348 (1959). He has calculated the energies of the ferromagnetic ground state and the excited states obtained from the ground state by reversing only one spin.

[16] Also see W. J. Carr, Phys. Rev. **92**, 28 (1953).

[17] The overlap matrix **D** in this paper is defined by (2.4). Here the ij element of **D** is given by the overlap between N-electron wave functions ψ_i and ψ_j. Under this definition of **D**, neither the inverse of **D** nor **D** itself has ever been calculated correctly when N is large. When the overlap matrix Δ is defined by overlaps between atomic orbitals so that its kh element is given by $\langle k|h\rangle$, the inverse Δ^{-1} has been expanded into a power series of Δ by P. O. Löwdin, Advan. Phys. **5**, 1 (1956). However, there is very little relation between Δ and **D**.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ground state carrying a suitable spin function Θ, the wave function is simply written as (2.1) and the calculation of $\mathbf{A}$ can be bypassed. This will simplify the calculation considerably as we shall describe in Sec. IV.

Most of the present discussion will be carried out as if we are dealing with arrays of one-electron atoms. However, this limitation is merely an aid to keep track of the formulas and manipulations. In Sec. IV(3), we shall give a brief discussion of problems which arise when arrays of many-electron atoms are considered.

Before we start to calculate $\mathbf{F}$ and $\mathbf{A}$ in Sec. III, we shall review the method used in the previous paper I in Sec. II.

II. THE HEITLER–LONDON METHOD

Let us consider an array of N one-electron atoms. The Heitler–London wave function for this system is given by

$$\Psi=(N!)^{-1/2}\sum_P^{N!}(-1)^P P^r\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)\times P^\sigma\Theta(\sigma_1,\sigma_2,\cdots,\sigma_N), \quad (2.1)$$

where $\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)$ is a product of occupied atomic orbitals $\chi_h(\mathbf{r}_h)$,

$$\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)=\chi_1(\mathbf{r}_1)\chi_2(\mathbf{r}_2)\cdots\chi_N(\mathbf{r}_N), \quad (2.2)$$

$\Theta(\sigma_1,\sigma_2,\cdots,\sigma_N)$ is an arbitrary function of the spin variables, and P^r and P^σ are permutations P of coordinate and spin variables. (P^r and P^σ are considered as corresponding to the same abstract permutation P.) The summation $\sum_P$ in (2.1) includes all $N!$ permutations P.

There are many ways to construct $\Theta(\sigma_1,\sigma_2,\cdots,\sigma_N)$ of N spins, but, in any case, we will find 2^N linearly independent and orthonormal functions $\Theta_1, \Theta_2, \cdots$. Starting from an orbital product $\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)$, therefore it is possible to generate, in principle, 2^N linearly independent functions $\Psi_1, \Psi_2, \cdots$ of this array. Let $\boldsymbol{\Theta}$ be the row matrix $\boldsymbol{\Theta}=(\Theta_1,\Theta_2,\cdots)$, whose elements are 2^N linearly independent and orthonormal functions Θ_i of N spins. Similarly $\boldsymbol{\Psi}$ is the row matrix $\boldsymbol{\Psi}=(\Psi_1,\Psi_2,\cdots)$. Then the wave functions Ψ_i introduced by (2.1) are written as

$$\boldsymbol{\Psi}=(N!)^{-1/2}\sum_P^{N!}(-1)^P P^r\Lambda P^\sigma\boldsymbol{\Theta} \quad (2.3)$$

by using the matrix representation.

The wave functions $\boldsymbol{\Psi}=(\Psi_1,\Psi_2,\cdots)$ thus obtained are not normalized because of the nonorthogonality of the atomic orbitals χ_k used as basis. Instead, it will be found that the overlap between Ψ_i and Ψ_j is given by

$$D_{ij}\equiv\int\Psi_i^*\Psi_j d\tau=\int[\sum_P(-1)^P P^r\Lambda P^\sigma\Theta_i]^*\Lambda\Theta_j d\tau=\sum_P\langle P\Lambda|\Lambda\rangle\tilde{U}_{ij}(P), \quad (2.4)$$

where $\langle P\Lambda|\Lambda\rangle$ represents the integral over the coordinate variables

$$\langle P\Lambda|\Lambda\rangle\equiv\int P^r\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)^*\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)dv, \quad (2.5)$$

while the integral over the spin variables is denoted by $\tilde{U}_{ij}(P)$ as follows:

$$\tilde{U}_{ij}(P)\equiv(-1)^P\int P^\sigma\Theta_i^*\Theta_j d\sigma. \quad (2.6)$$

From (2.4), it is clear that the normalization constants D_{ii} can become very large, which leads to the difficulty of calculating the Heitler–London energy.

The overlap integral D_{ij} can be regarded as the ij element of the overlap matrix $\mathbf{D}$, and similarly $\tilde{U}_{ij}(P)$ as the ij element of matrix $\tilde{\mathbf{U}}(P)$. Using the matrix notation, Eqs. (2.4) and (2.6) are written as

$$\mathbf{D}\equiv\int\boldsymbol{\Psi}^\dagger\boldsymbol{\Psi}d\tau=\sum_P\langle P\Lambda|\Lambda\rangle\tilde{\mathbf{U}}(P), \quad (2.7)$$

and

$$\tilde{\mathbf{U}}(P)=(-1)^P\int P^\sigma\boldsymbol{\Theta}^\dagger\cdot\boldsymbol{\Theta}d\sigma. \quad (2.8)$$

It is easily proved that the matrix $\tilde{\mathbf{U}}(PQ)$ corresponding to the successive application of two permutations Q and P is given by the product of two matrices $\tilde{\mathbf{U}}(Q)$ and $\tilde{\mathbf{U}}(P)$ as follows[18,19]:

$$\tilde{\mathbf{U}}(PQ)=\tilde{\mathbf{U}}(Q)\tilde{\mathbf{U}}(P). \quad (2.9)$$

---

[18] The matrix $\mathbf{U}(P)$ introduced by (2.8) is unitary and is a representation matrix of the permutation group. Use of representation matrices of permutation groups in the calculation of electronic energy levels is originally due to R. Serber [Phys. Rev. **45**, 461 (1934)] and T. Yamanouchi [Proc. Phys. Math. Soc. Japan **18**, 623 (1936)]. Also see the review by M. Kotani in *Table of Molecular Integrals*, edited by M. Kotani *et al.* (Maruzen Co., Ltd., Tokyo, 1955). In the present paper, however, we shall not use any group-theoretical argument except in (2.9) and (3.39). It is more convenient to understand that $\tilde{\mathbf{U}}(P)$ is merely a notation representing the integral appearing on the right of (2.8) and has the property described by (2.9).

[19] The relation (2.9) will be obtained as follows. Since 2^N linearly independent and orthonormal spin functions Θ_i, for $i=1, 2, \cdots, 2^N$, are complete in the spin space, we can expand $P^\sigma\Theta_i$ in terms of Θ_i as follows:

$$P^\sigma\Theta_i=(-1)^P\sum_j\Theta_j U_{ji}(P)^*, \quad (1)$$

where $U_{ji}(P)^*$ is merely an expansion coefficient but it is easily recognized that $U_{ji}(P)$ defined above is equivalent to $\tilde{U}_{ij}(P)$ introduced by (2.6). Here $\tilde{\mathbf{U}}(P)$ is the transpose of matrix $\mathbf{U}(P)$. Let us apply two permutations Q^σ and P^σ successively to Θ_i. Then

$$P^\sigma Q^\sigma\Theta_i=(-1)^Q\sum_k P^\sigma\Theta_k U_{ki}(Q)^*=(-1)^{PQ}\sum_j\sum_k\Theta_j U_{jk}(P)^*U_{ki}(Q)^*. \quad (2)$$

On the other hand, we can write this as follows:

$$P^\sigma Q^\sigma\Theta_i=(-1)^{PQ}\sum_j\Theta_j U_{ji}(PQ)^*. \quad (3)$$

Comparison of the above two equations yields

$$\mathbf{U}(PQ)=\mathbf{U}(P)\,\mathbf{U}(Q), \quad (4)$$

which is equivalent to (2.9).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The orthonormal wave functions Ψ_N of the array are then given by

$$\Psi_N=\Psi\mathbf{D}^{-1/2}. \tag{2.10}$$

Our problem is to calculate the Heitler–London energy of a Hamiltonian of the form

$$\mathcal{H}=\sum_{h=1}^{N} f(\mathbf{r}_h)+\tfrac{1}{2}\sum_{h_1=1}^{N}\sum_{h_2=1}^{N}{}' g(\mathbf{r}_{h_1},\mathbf{r}_{h_2}), \tag{2.11}$$

where the summation $\sum_{h_2=1}^{N}{}'$ excludes $h_2=h_1$. For this purpose, it is convenient to split off, from the N-electron wave function Ψ, one or several atomic orbitals, which are of particular interest for the manipulations we want to perform. This will be carried out as follows. An antisymmetric wave function of N electrons is constructed by taking the summation over $N!$ permutations P as is shown in (2.3). Here the $N!$ permutations P can be divided into two sets; the first set includes $(N-1)!$ permutations $Q[h]$ of $N-1$ electrons $1, 2, \cdots, h-1, h+1, \cdots, N$, while the second one consists of N permutations $P(k\leftarrow h)$ taking electron h to orbital k, where $k=1, 2, \cdots, N$. The original $N!$ permutations are given by $Q[h]\cdot P(k\leftarrow h)$ and the summation over $N!$ permutations P in (2.3) splits into two; the summations over $(N-1)!$ permutations $Q[h]$ and over k's. If we take the summation over $(N-1)!$ permutations $Q[h]$ after operating a particular permutation $P(k\leftarrow h)$, this portion of the wave function is antisymmetric with respect to electrons $1, 2, \cdots, h-1, h+1, \cdots, N$, but electron h always occupies orbital k. To emphasize this situation, let us denote the sum by $\chi_k(\mathbf{r}_h)\Psi_i[k(h)]$ where $i=1, 2, \cdots, N$. The original wave function Ψ is then written as

$$\Psi=N^{-1/2}\sum_k \chi_k(\mathbf{r}_h)\Psi[k(h)], \tag{2.12}$$

where $\Psi[k(h)]$ is the row matrix whose elements $\Psi_i[k(h)]$ are functions of $N-1$ orbitals and N spins as described above. More explicitly,

$$\Psi[k(h)]\equiv[(N-1)!]^{-1/2}\sum_{Q[h]}(-1)^{Q[h]P(k\leftarrow h)} \times Q^r[h]\Lambda[k(h)]Q^\sigma[h]P^\sigma(k\leftarrow h)\Theta, \tag{2.13}$$

and $\Lambda[k(h)]$ is a product of $N-1$ atomic orbitals, obtained from $\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)$ by applying $P^r(k\leftarrow h)$ and deleting $\chi_k(\mathbf{r}_h)$ as follows:

$$\Lambda[k(h)]\equiv\frac{1}{\chi_k(\mathbf{r}_h)}P^r(k\leftarrow h)\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N). \tag{2.14}$$

Similarly we introduce

$$\Psi[k_1(h_1)k_2(h_2)],\ \cdots,\ \Psi[k_1(h_1)\cdots k_n(h_n)].$$

For example,

$$\Psi[k_1(h_1)k_2(h_2)] \equiv[(N-2)!]^{-1/2}\sum_{Q[h_1h_2]}(-1)^{Q[h_1h_2]P(k_1k_2\leftarrow h_1h_2)} \times Q^r[h_1h_2]\Lambda[k_1(h_1)k_2(h_2)]Q^\sigma[h_1h_2] \times P^\sigma(k_1k_2\leftarrow h_1h_2)\Theta, \tag{2.15}$$

where

$$\Lambda[k_1(h_1)k_2(h_2)] \equiv\frac{1}{\chi_{k_1}(\mathbf{r}_{h_1})\chi_{k_2}(\mathbf{r}_{h_2})}P^\sigma(k_1k_2\leftarrow h_1h_2) \times\Lambda(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N), \tag{2.16}$$

and $Q[h_1h_2]$ represents $(N-2)!$ permutations not involving electrons h_1 and h_2, while $P(k_1k_2\leftarrow h_1h_2)$ is a permutation taking electrons h_1 and h_2 to orbitals k_1 and k_2. The functions

$$\Psi[k_1(h_1)k_2(h_2)],\quad \Psi[k_1(h_1)k_2(h_2)k_3(h_3)],$$

etc., are related to the original wave functions Ψ as follows:

$$\Psi=[N(N-1)]^{-1/2}\sum_{k_1}\sum_{k_2}{}'\chi_{k_1}(\mathbf{r}_{h_1})\chi_{k_2}(\mathbf{r}_{h_2}) \times\Psi[k_1(h_1)k_2(h_2)]=\text{etc.} \tag{2.17}$$

We further define the overlap matrices

$$S[k_1\cdots k_n|h_1\cdots h_n]$$

between the functions $\Psi[k_1(h_1)\cdots k_n(h_n)]$ as follows:

$$\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]\equiv\int\Psi[k_1(h_1)\cdots k_n(h_n)]^\dagger \times\Psi[k_1(h_1)\cdots k_n(h_n)]d\tau. \tag{2.18}$$

The method used in obtaining (2.4) is applied to the above matrices, which leads to

$$\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]=\sum_{Q[h_1\cdots h_n]}\langle Q[h_1\cdots h_n]\Lambda[k_1(h_1)\cdots k_n(h_n)|\Lambda[h_1(h_1)\cdots h_n(h_n)]\rangle \times\tilde{\mathbf{U}}(Q[h_1\cdots h_n]P(k_1\cdots k_n\leftarrow h_1\cdots h_n)). \tag{2.19}$$

Those overlap matrices are again related to the overlap matrix $\mathbf{D}$ between the original wave functions Ψ intro-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

duced by (2.4) as follows:

$$\mathbf{D}=\mathbf{S}[h|h]+\sum_k{}'\langle k|h\rangle\mathbf{S}[k|h] \tag{2.20}$$

$$=\mathbf{S}[h_1h_2|h_1h_2]+\langle h_1|h_2\rangle\langle h_2|h_1\rangle\mathbf{S}[h_2h_1|h_1h_2]$$
$$+\sum_k{}'\{\langle k|h_1\rangle\mathbf{S}[kh_2|h_1h_2]+\langle k|h_2\rangle\mathbf{S}[h_1k|h_1h_2]+\langle k|h_2\rangle\langle h_2|h_1\rangle\mathbf{S}[h_2k|h_1h_2]+\langle k|h_1\rangle\langle h_1|h_2\rangle\mathbf{S}[kh_1|h_1h_2]\}$$
$$+\sum_{k_1k_2}{}''\langle k_1|h_1\rangle\langle k_2|h_2\rangle\mathbf{S}[k_1k_2|h_1h_2] \tag{2.21}$$

$$=\text{etc.},$$

where

$$\langle k|h\rangle=\int\chi_k(\mathbf{r}_h)^*\chi_h(\mathbf{r}_h)d\tau_h, \tag{2.22}$$

and $\sum_k{}'$ excludes $k=h$, and $\sum_{k_1k_2}{}''$ excludes $k_1=h_1$, $k_2=h_2$ and also $k_1=h_2$, $k_2=h_1$. These results can be proved by inserting (2.19) into (2.20) and (2.21).

More generally, it will be found that

$$\mathbf{S}[k_1\cdots k_{n-1}|h_1\cdots h_{n-1}]$$
$$=\delta(h_n\neq k_1,\cdots,k_{n-1})\mathbf{S}[k_1\cdots k_{n-1}h_n|h_1\cdots h_{n-1}h_n]$$
$$+\sum_{k_n}{}^{(n)}\langle k_n|h_n\rangle\mathbf{S}[k_1\cdots k_{n-1}k_n|h_1\cdots h_{n-1}h_n], \tag{2.23}$$

where

$$\delta(h_n\neq k_1,\cdots,k_{n-1})=1, \quad \text{if } h_n \text{ is not included in } k_1,\cdots,k_{n-1},$$
$$=0, \quad \text{otherwise}$$

and, in the summation $\sum_{k_n}{}^{(n)}$, we exclude $k_n=k_1, k_2, \cdots, k_{n-1}, h_n$. Furthermore, the following equality will hold:

$$\mathbf{S}[h_{i_1}h_{i_2}\cdots h_{i_n}|h_1h_2\cdots h_n]=\tilde{\mathbf{U}}(P_i)\mathbf{S}[h_1h_2\cdots h_n|h_1h_2\cdots h_n], \tag{2.24}$$

where P_i is a permutation in which $h_1h_2\cdots h_n$ are taken to orbitals $h_{i_1}h_{i_2}\cdots h_{i_n}$. This relation will be proved easily by replacing $P(k_1\cdots k_n\leftarrow h_1\cdots h_n)$ in (2.19) by P_i and by splitting matrices $\tilde{\mathbf{U}}(Q[h_1\cdots h_n]P_i)$ as $\tilde{\mathbf{U}}(P_i)\tilde{\mathbf{U}}(Q[h_1\cdots h_n])$ by using the relation (2.9).

Now we are ready to calculate the energy matrix of the Hamiltonian (2.11). Since the Hamiltonian is totally symmetric and commutes with the antisymmetrizer $\sum_P(-1)^PP$ in the wave functions (2.3), the method used in obtaining (2.4) is again applied to this calculation and it will be found that

$$\langle\Psi|\sum_h f(\mathbf{r}_h)|\Psi\rangle\equiv\int\Psi^\dagger\sum_h f(\mathbf{r}_h)\Psi d\tau$$
$$=\sum_h\langle h|f(\mathbf{r}_h)|h\rangle\mathbf{S}[h|h]$$
$$+\sum_h\sum_k{}'\langle k|f(\mathbf{r}_h)|h\rangle\mathbf{S}[k|h]. \tag{2.25}$$

The normalized expectation values $\langle\mathbf{f}\rangle$, calculated by using the orthonormal wave functions Ψ_N defined by (2.10), are given by

$$\langle\mathbf{f}\rangle\equiv\langle\Psi_N|\sum_h f(\mathbf{r}_h)|\Psi_N\rangle$$
$$=\sum_h\langle h|f(\mathbf{r}_h)|h\rangle\mathbf{D}^{-1/2}\mathbf{S}[h|h]\mathbf{D}^{-1/2}$$
$$+\sum_h\sum_k{}'\langle k|f(\mathbf{r}_h)|h\rangle\mathbf{D}^{-1/2}\mathbf{S}[k|h]\mathbf{D}^{-1/2}. \tag{2.26}$$

Similarly the expectation values $\langle\mathbf{g}\rangle$ of two-body interaction $\frac{1}{2}\sum_{h_1}\sum_{h_2}{}'g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})$ are obtained by

$$\langle\mathbf{g}\rangle\equiv\langle\Phi_N|\tfrac{1}{2}\sum_{h_1}\sum_{h_2}{}'g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|\Phi_N\rangle$$
$$=\tfrac{1}{2}\sum_{h_1}\sum_{h_2}\sum_{k_1}\sum_{k_2}\langle k_1k_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$$
$$\times\mathbf{D}^{-1/2}\mathbf{S}[k_1k_2|h_1h_2]\mathbf{D}^{-1/2}, \tag{2.27}$$

where h_1 and h_2 are also included in the summations $\sum_{k_1}\sum_{k_2}{}'$ for simplicity.

Here $\langle\mathbf{f}\rangle$ and $\langle\mathbf{g}\rangle$ are both matrices and can be calculated if the normalization matrices $\mathbf{D}^{-1/2}\mathbf{S}[k|h]\mathbf{D}^{-1/2}$ and $\mathbf{D}^{-1/2}\mathbf{S}[k_1k_2|h_1h_2]\mathbf{D}^{-1/2}$ are evaluated explicitly. The calculation of the normalization matrices is the main subject of the present paper and will be carried out in the following section.

III. CALCULATION OF THE NORMALIZATION MATRICES

Here we shall expand the normalization matrices of the type $\mathbf{D}^{-1/2}\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]\mathbf{D}^{-1/2}$ in terms of $\tilde{\mathbf{U}}(P)$ and evaluate the expansion coefficients. To proceed with the problem, let us divide the expansion into two steps and first consider the calculation of matrices of the following form:

$$\mathbf{F}[k_1\cdots k_n|h_1\cdots h_n]\equiv\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]\mathbf{D}^{-1}$$
$$=\sum_P f_P[k_1\cdots k_n|h_1\cdots h_n]\tilde{\mathbf{U}}(P). \tag{3.1}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The second expansion to be calculated is of the form:

$$\mathbf{A}^{(1/2)}(R)\equiv\mathbf{D}^{-1/2}\tilde{\mathbf{U}}(R)\mathbf{D}^{1/2}=\sum_P a_P{}^{(1/2)}(R)\tilde{\mathbf{U}}(P). \quad (3.2)$$

The original matrix is thus expanded as

$$\begin{aligned}&\mathbf{D}^{-1/2}\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]\mathbf{D}^{-1/2}\\&=\mathbf{D}^{-1/2}\mathbf{S}[k_1\cdots k_n|h_1\cdots h_n]\mathbf{D}^{-1}\cdot\mathbf{D}^{1/2},\\&=\sum_Q f_Q[k_1\cdots k_n|h_1\cdots h_n]\mathbf{D}^{-1/2}\tilde{\mathbf{U}}(Q)\mathbf{D}^{1/2},\\&=\sum_P\sum_Q f_Q[k_1\cdots k_n|h_1\cdots h_n]a_P{}^{(1/2)}(Q)\tilde{\mathbf{U}}(P). \quad (3.3)\end{aligned}$$

Previously,[10,20] we have shown that, as long as atomic orbitals χ used as basis are linearly independent, the inverse $\mathbf{D}^{-1}$ exists and can be expanded as (24) of paper I, that is,

$$\mathbf{D}^{-1}=\sum_P T_P\tilde{\mathbf{U}}(P). \quad (3.4)$$

Multiplication of this expression with (2.19) yields (3.1). The second equation (3.2), which is similarly obtainable, has been introduced previously by (68) of paper I.

Our problem now is to expand the coefficients $f_P[k_1\cdots k_n|h_1\cdots h_n]$ and $a_P{}^{(1/2)}(R)$ in terms of overlap integrals so that f_P and $a_P{}^{(1/2)}(R)$ can be evaluated explicitly.

(1) Expansion of the F Matrices

(a) *The Recurrence Relations for the Expansion*

Let us first consider the matrices of the form $\mathbf{F}[h_1\cdots h_n|h_1\cdots h_n]$ and show that, by using the relations (2.20), (2.21), and (2.23), the matrices can be decomposed into a sum of products of two parts: The first ones denoted by $\Gamma[P]$ are made of overlaps between electrons $h_1\cdots h_n$ and the other electrons, and can be calculated up to any desired accuracy explicitly while the second parts are equivalent to the $\mathbf{F}$ matrices originally considered except that the new matrices are for a new array which can be obtained by removing electrons $h_1\cdots h_n$ from the original array. Therefore we can repeat the same calculation in the second parts. This expansion will converge since overlap between electron h and a distant electron decreases exponentially.

For example let us consider the simplest case, the matrix $\mathbf{F}[h|h]=\mathbf{S}[h|h]\mathbf{D}^{-1}$. By inserting (2.20) into $\mathbf{F}[h|h]$, we find that

$$\begin{aligned}\mathbf{F}[h|h]&=\mathbf{S}[h|h]\{\mathbf{S}[h|h]+\sum_k{}'\langle k|h\rangle\mathbf{S}[k|h]\}^{-1}\\&=(1+\mathbf{X}[h])^{-1},\end{aligned} \quad (3.5)$$

where

$$\mathbf{X}[h]\equiv\sum_k{}'\langle k|h\rangle\mathbf{S}[k|h]\mathbf{S}[h|h]^{-1}. \quad (3.6)$$

Repeated use of (2.23) in $\mathbf{X}[h]$ yields

$$\begin{aligned}\mathbf{X}[h]=&\sum_{k_1}{}'\langle h|k_1\rangle\langle k_1|h\rangle\mathbf{S}[k_1h|hk_1]\mathbf{S}[h|h]^{-1}\\&+\sum_{k_2}{}''\sum_{k_1}{}'\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\mathbf{S}[k_1k_2h|hk_1k_2]\mathbf{S}[h|h]^{-1}\\&+\sum_{k_3}{}'''\sum_{k_2}{}''\sum_{k_1}{}'\langle h|k_3\rangle\langle k_3|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\mathbf{S}[k_1k_2k_3h|hk_1k_2k_3]\mathbf{S}[h|h]^{-1}\\&+\cdots.\end{aligned} \quad (3.7)$$

Let us introduce the notation $\mathbf{F}_h[k_1\cdots k_n|k_1\cdots k_n]$ by defining

$$\mathbf{S}[hk_1\cdots k_n|hk_1\cdots k_n]\mathbf{S}[h|h]^{-1}\equiv\mathbf{F}_h[k_1\cdots k_n|k_1\cdots k_n]. \quad (3.8)$$

Use of this definition as well as (2.24) leads to

$$\begin{aligned}\mathbf{X}[h]=&\sum_{k_1}{}'\langle h|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1h)\mathbf{F}_h[k_1|k_1]\\&+\sum_{k_2}{}''\sum_{k_1}{}'\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1k_2h)\mathbf{F}_h[k_1k_2|k_1k_2]\\&+\sum_{k_3}{}'''\sum_{k_2}{}''\sum_{k_1}{}'\langle h|k_3\rangle\langle k_3|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1k_2k_3h)\mathbf{F}_h[k_1k_2k_3|k_1k_2k_3]\\&+\cdots,\end{aligned} \quad (3.9)$$

where $(k_1k_2\cdots k_nh)$ is a cyclic permutation in which electrons $h, k_1, \cdots, k_{n-1}, k_n$ are shifted to orbitals $k_1, k_2, \cdots, k_n, h$. In (3.9), terms such as $\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1k_2h)$ belong to the first part $\Gamma[P]$, while the matrices $\mathbf{F}_h[k_1\cdots k_n|k_1\cdots k_n]$ belonging to the second part are equivalent to $\mathbf{F}[k_1\cdots k_n|k_1\cdots k_n]$. For instance, $\mathbf{F}_h[k|k]$ can be calculated by (3.5) and (3.9) except that electron h should be excluded from the summations.

[20] The present method can be extended to the case where basic atomic orbitals are not linearly independent. See Sec. IV (3) and also Appendix B of paper I.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The equation (3.5) can be expanded formally as

$$\mathbf{F}[h|h]=1-\mathbf{X}[h]+\mathbf{X}[h]^2-\mathbf{X}[h]^3+\cdots. \tag{3.10}$$

This expansion converges as long as $\epsilon=\sum_P|x_P[h]|<1$ where $x_P[h]$ is the expansion coefficient of $\mathbf{X}[h]$:

$$\mathbf{X}[h]=\sum_P x_P[h]\tilde{\mathbf{U}}(P). \tag{3.11}$$

This will be proved in the Appendix.

Similarly $\mathbf{F}[h_1h_2|h_1h_2]$ is expanded as

$$\begin{aligned}\mathbf{F}[h_1h_2|h_1h_2]&=(1+\mathbf{X}[h_1h_2])^{-1}\\&\doteq 1-\mathbf{X}[h_1h_2]+\mathbf{X}[h_1h_2]^2-\mathbf{X}[h_1h_2]^3+\cdots,\end{aligned} \tag{3.12}$$

where

$$\begin{aligned}\mathbf{X}[h_1h_2]=&\langle h_1|h_2\rangle\langle h_2|h_1\rangle\tilde{\mathbf{U}}(h_1h_2)\\&+\sum_{k_1}\{\langle h_1|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(k_1h_1)+\langle h_2|k_1\rangle\langle k_1|h_2\rangle\mathbf{U}(k_1h_2)+\langle h_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|h_2\rangle\tilde{\mathbf{U}}(k_1h_2h_1)\\&\qquad+\langle h_1|k_1\rangle\langle k_1|h_2\rangle\langle h_2|h_1\rangle\tilde{\mathbf{U}}(h_2k_1h_1)\}\mathbf{F}_{h_1h_2}[k_1|k_1]\\&+\sum_{k_1k_2}\{\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(k_1k_2h_1)+\langle h_2|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_2\rangle\tilde{\mathbf{U}}(k_1k_2h_2)\\&\quad+\langle h_2|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|h_2\rangle\tilde{\mathbf{U}}(k_1k_2h_2h_1)+\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_2\rangle\langle h_2|h_1\rangle\tilde{\mathbf{U}}(h_2k_1k_2h_1)\\&\quad+\langle h_1|k_1\rangle\langle k_1|h_1\rangle\langle h_2|k_2\rangle\langle k_2|h_2\rangle\tilde{\mathbf{U}}((k_1h_1)(k_2h_2))\\&\quad+\langle h_1|k_2\rangle\langle k_2|h_2\rangle\langle h_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(k_1h_2k_2h_1)\}\mathbf{F}_{h_1h_2}[k_1k_2|k_1k_2]\\&+\cdots.\end{aligned} \tag{3.13}$$

The matrices $\mathbf{F}[h_1|h_1]$, $\mathbf{F}[h_1h_2|h_1h_2]$, $\cdots$, obtained by (3.9), (3.10), (3.12), and (3.13), can be inserted into replace (3.9) and (3.13). Thus the expansion develops further. These relations are more conveniently summarized by

$$\begin{aligned}\mathbf{F}[h(n)]&=(1+\mathbf{X}[h(n)])^{-1}\\&=1-\mathbf{X}[h(n)]+\mathbf{X}[h(n)]^2-\mathbf{X}[h(n)]^3+\cdots,\end{aligned} \tag{3.14}$$

and

$$\mathbf{X}[h(n)]=\sum_m\sum_k\Gamma[k(m)]\mathbf{F}_{h(n)}[k(m)], \tag{3.15}$$

where $h(n)$ and $k(n)$ are abbreviations of letters appearing in brackets of the matrices $\mathbf{F}[h_1\cdots h_n|h_1\cdots h_n]$, $\mathbf{X}[h_1\cdots h_n]$, etc., and n specifies the number of letters involved.

So far we consider matrices of the form $\mathbf{F}[h_1\cdots h_n|h_1\cdots h_n]$. More general ones $\mathbf{F}[k_1\cdots k_n|h_1\cdots h_n]$ introduced by (3.1) can also be reduced to the standard form by repeated use of (2.23). For instance,

$$\begin{aligned}\mathbf{F}[k|h]=&\langle h|k\rangle\tilde{\mathbf{U}}(hk)\mathbf{F}[hk|hk]\\&+\sum_{k_1}{}'\langle h|k_1\rangle\langle k_1|k\rangle\tilde{\mathbf{U}}(kk_1h)\mathbf{F}[hkk_1|hkk_1]\\&+\sum_{k_2}{}''\sum_{k_1}{}'\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|k\rangle\tilde{\mathbf{U}}(kk_1k_2h)\mathbf{F}[hkk_1k_2|hkk_1k_2]\\&+\cdots.\end{aligned} \tag{3.16}$$

(b) *Diagram Description*

To proceed with the calculation in practice, it is convenient to draw diagrams. Let us denote, by a solid line starting from point h_1 and ending at point h_2, the overlap integral $\langle h_2|h_1\rangle$ and similarly, by a broken line starting from h_1 going through $h_2,\cdots,h_n$ and coming back to h_1, a cyclic permutation $(h_2h_3\cdots h_nh_1)$. For example, the expansion (3.9) of $\mathbf{X}[h]$ is shown in Fig. 1, where, as a convenient way of explaining how diagrams are built up, points are classified *temporarily* as the following two kinds: a solid dot ● indicates the starting point of $\mathbf{X}[h]$ and no line comes back to this point any more in the course of expanding $\mathbf{X}[h]$, while an open dot ○ indicates that the expansion is not completed at

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

this point but a new $\mathbf{F}$ matrix starts from here. The first term has one open dot and therefore $\mathbf{F}_h[k_1|k_1]$ will follow. Two open dots of the second term show that $\mathbf{F}_h[k_1k_2|k_1k_2]$ follows. Since broken lines for permutations are the same as the solid lines, we have omitted them.

Similarly the expansion (3.13) of $\mathbf{X}[h_1h_2]$ is described in Fig. 2. There we have omitted terms obtained by interchanging h_1 and h_2.

To calculate $\mathbf{F}_h[k_1|k_1]$, $\mathbf{F}_h[k_1k_2|k_1k_2]$, etc., we further draw $\mathbf{X}_h[k_1|k_1]$, $\mathbf{X}_h[k_1k_2|k_1k_2]$, etc., starting from open dots. For example, the second term in Fig. 1 includes the series of diagrams in Fig. 3.

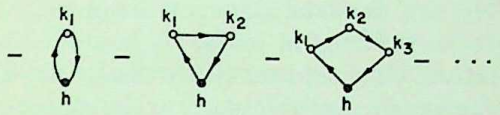


FIG. 1. The expansion of $\mathbf{X}[h]$ described by Eq. (3.9).

Solid lines can come into a point many times, but the representation of electron permutations should be reduced into the simplest form so that we can distinguish a permutation from others immediately. This implies that within one diagram, any one electron should not be shifted more than once, and hence a broken line can come into a point only once and come out only once. For instance, the second term in Fig. 3 represents the permutation $(k_1k_2)(k_1k_2h_1)$ which shifts electrons h_1, k_1, k_2 to k_1, k_2, h_1 and, after that, again interchanges k_1 and k_2, but this process is equivalent to a single interchange

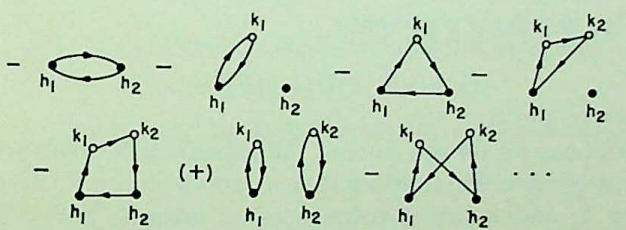


FIG. 2. The expansion of $\mathbf{X}[h_1h_2]$ described by Eq. (3.13).

of h_1 and k_2. Therefore the broken loop in this diagram illustrates the interchange (h_1k_2) only. The second term represents the quantity

$$|\langle k_1|k_2\rangle|^2\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_2).$$

Since $(k_1k_2h_1)(k_1k_2)=(h_1k_1)$ is a different permutation from the second term, it is important to remember the order in which these loops appear in a diagram.

In order to calculate $\mathbf{F}[h|h]$ by (3.14), we need to evaluate higher order terms $\mathbf{X}[h]^2$, $\mathbf{X}[h]^3$, $\cdots$. This will be carried out by multiplying diagrams we have generated. For instance, $\mathbf{X}[h]^2$ will be obtained from diagrams in Fig. 1. In case permutation $(k_1k_2h_1)$ appears first (on the left of $\mathbf{X}[h]\mathbf{X}[h]$) and then permutation (h_1k_1) follows it [(h_1k_1) appears on the right of $\mathbf{X}[h]\mathbf{X}[h]$], we obtain the diagram in Fig. 4, which corresponds to the quantity

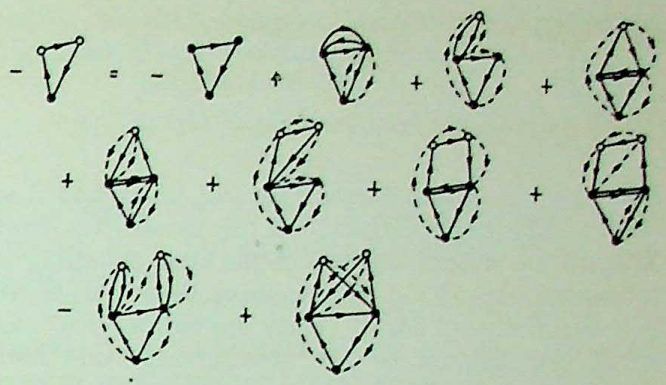


FIG. 3. Further expansion of the second term in Fig. 1.

$$|\langle h|k_1\rangle|^2\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1k_2).$$

Since all diagrams in $\mathbf{X}[h]$ start from h, solid lines come into h many times after the multiplication and the distinction between solid and open dots becomes obscure. Therefore, we shall not make this distinction in diagrams any more.

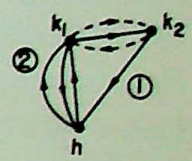


FIG. 4. The diagram of $\langle h|k_2\rangle\langle k_2|k_1\rangle \times\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1k_2h)\times\langle h|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1h)$. This term comes from $\mathbf{X}[h]^2$.

Similar multiplication should be added in those appearing in Fig. 3. Here we have drawn only $\mathbf{X}_h[k_1k_2|k_1k_2]$ instead of $\mathbf{F}_h[k_1k_2|k_1k_2]$. The expression will become complete if we include higher terms such as $\mathbf{X}_h[k_1k_2|k_1k_2]^2$, $\mathbf{X}_h[k_1k_2|k_1k_2]^3$, etc. The contribution from $\mathbf{X}_h[k_1k_2|k_1k_2]^2$ includes, for instance, the diagrams in Fig. 5. The first one is obtained by multiplication of the first and second diagrams in Fig. 2. The second one is obtained from two of the type appearing in the second term, but h_1 and h_2 are interchanged in the first one.

In (3.14), odd powered terms $\mathbf{X}[h(n)]^{2m+1}$ carry minus signs, which should also appear in diagrams. It is possible to associate a sign with the number of loops in a diagram. For instance, diagrams in Fig. 1 should have minus signs in $\mathbf{F}[h_1|h_1]$ since they are terms appearing in $\mathbf{X}[h_1]$. In fact, each of them consists of one loop. Those in Fig. 2 should also be minus. However, the fifth

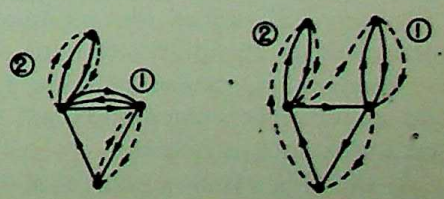


FIG. 5. Some of the contributing terms from $\mathbf{X}_h[k_1k_2|k_1k_2]^2$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

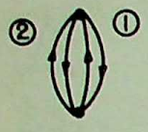


FIG. 6. The diagram of $\langle h|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1h)\times\langle h|k_1\rangle\langle k_1|h\rangle\tilde{\mathbf{U}}(k_1h)$.

term consists of two loops. This will be explained as follows. The same diagram can be constructed from $\mathbf{X}[h_1h_2]^2$ by multiplying two of the type appearing in the second term in Fig. 2, as we have already explained in connection with Fig. 5. There are two ways to construct the diagram by the multiplication and both carry a plus sign. Therefore, the total sum is $2-1=1$ in accordance with the proposed criterion. Matrix $\mathbf{X}[h_1h_2h_3]$ will include diagrams with three loops. However, the same diagram will be constructed from $\mathbf{X}[h_1h_2h_3]^2$ 6 times with plus sign and from $\mathbf{X}[h_1h_2h_3]^3$ 6 times with minus sign. The total is $-1+6-6=(-1)^3$. We can test this rule further. Let us assume that, up to t loops, total sum of all possible contributions to a particular diagram with r loops is $(-1)^r$, where $r\leqslant t$. A diagram with $t+1$ loops is constructed by multiplying a diagram with χ loops by all possible diagrams with total $t-\chi+1$ loops. By considering all possible ways to construct diagrams each of which carries χ loops and by adding all possible cases starting from $\chi=1$ and counting up to $\chi=t+1$, we obtain the total sum of all possible contributions to the diagram with $t+1$ loops. We can construct $\binom{t+1}{\chi}$ diagrams, each of which carries χ loops, while the total sum of all possible contributions to a diagram with $t-\chi+1$ loops is $(-1)^{t-\chi+1}$. Therefore, the total sum of all possible contributions to the desired diagram carrying $t+1$ loops is

$$\sum_{\chi=1}^{t+1}(-1)^{t-\chi+1}\binom{t+1}{\chi}=(-1)^{t+1}.$$

This proves the criterion on assigning signs to diagrams: *We count the number t of loops in solid lines in a diagram. The diagram should carry sign* $(-1)^t$.

Thus we have expanded the matrix $\mathbf{F}[h_1\cdots h_n|h_1\cdots h_n]$ in terms of overlap integrals and $\tilde{\mathbf{U}}(P)$. Comparison of the expansion with (3.1) yields that

$$f_P[h_1\cdots h_n|h_1\cdots h_n]=\sum_{t=1}^{\infty}\sum_{P_1P_2\cdots P_t}^{P}(-1)^tS_{P_1}S_{P_2}\cdots S_{P_t},\quad(3.17)$$

where P_i is a cyclic permutation and $S_{P_i}=\int P_i\Lambda^*\Lambda dv$ represents a loop. The summation takes all possible $P_1, P_2, \cdots, P_t$ under the condition $P=P_tP_{t-1}\cdots P_1$ and, at least, P_1 will include some of $h_1h_2\cdots h_n$.

From (3.17), we can calculate $f_P[h_1\cdots h_n|h_1\cdots h_n]$ for any P. The computation will be simplified if we draw diagrams as follows:

(i) We start from a point in $h_1h_2\cdots h_n$, and draw a loop. A loop must never go through a point more than once and always come back to the starting point.

(ii) By adding more loops, we can calculate terms in high order. Loops appear by successive expansion (Fig. 3) of $\mathbf{F}$ and by multiplication (Figs. 4 and 5) of $\mathbf{X}$. Number the loops according to the order with which the loops appear and attach sign $(-1)^t$ to the diagram, where t is the number of loops in the diagram.

(iii) After completing solid lines in a diagram, we describe the electron permutation by broken lines. The broken lines never go through one point more than once in a diagram and always come back to the starting point. If the line is equal to P, retain the diagram. Otherwise discard it.

(iv) We can draw all different diagrams, including differences in ordering of loops. If, however, loops are commutative, we should take only one of the diagrams we find by numbering the loops in different ways. An example of such a case is the second diagram in Fig. 5, where we have already counted all contributions when $(-1)^t$ was assigned to it. Another example is shown in Fig. 6, which, in fact, appears only once in the expansion.

(c) *Accuracy of the Expansion Technique*

In practice, the expansion described in this section cannot be carried out indefinitely and it is necessary to truncate the series. The upper bound to an error introduced by such truncation of the series can be calculated by using the inequality (A7) for sum L_n of absolute values of the expansion coefficients $f_P[h(n)]$.

Let us consider, for example, the series in Fig. 1. The first term represents

$$\langle h|k_1\rangle\langle k_1|h\rangle\mathbf{U}(k_1h)\mathbf{F}_h[k_1|k_1].$$

Suppose we do not pursue the expansion beyond point k_1 any more. This implies that we approximate $\mathbf{F}_h[k_1|k_1]$ by 1, and hence it introduces an error to the series. From (A7), it is easily found that the error is less than

$$\sum_{k_1}\langle h|k_1\rangle\langle k_1|h\rangle[(1-O(\Delta\Delta_0))^{-1}-1].$$

From the second term in Fig. 1, we get an error of

$$\sum_{k_1}\sum_{k_2}\langle h|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h\rangle[(1-O(\Delta\Delta_0))^{-2}-1],$$

by not taking further expansion terms beyond k_1 and k_2, and so on. Here the factor $[(1-O(\Delta\Delta_0))^{-m}-1]$ will increase when m increases, and the first term $(1-O(\Delta\Delta_0))^{-m}$ will eventually become dominant. Under such circumstances, there is no reason to write the large loop without further expansion, since the loop is merely a minority term in the further expansion. They should be neglected altogether. For instance, if we neglect, in

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the series in Fig. 1, the fourth term and higher, the error will be

$$\Delta^5(1-O(\Delta\Delta_0))^{-4}+\Delta^6(1-O(\Delta\Delta_0))^{-5}+\Delta^7(1-O(\Delta\Delta_0))^{-6}+\cdots.$$

A similar observation can be made for the series in Fig. 3. If we retain the first term and neglect all the others, the error will be

$$\Delta^3[(1-O(\Delta\Delta_0))^{-2}-1].$$

If we consider all diagrams that appear in Fig. 3 and neglect the higher orders, we have to estimate the error for each diagram one by one. This process is not so difficult. The first two diagrams are complete and therefore introduce no error. The next three carry an error $[(1-O(\Delta\Delta_0))^{-1}-1]$ times their values and so on. In addition, an error will be introduced by not considering terms such as $X_h[k(2)]^2$, $X_h[k(2)]^3$, etc.

In any case, it is possible to calculate the upper bound to an error introduced by truncation of an expansion of $\mathbf{F}$.

(2) Calculation of the A Matrices

(a) *Evaluation of* $\mathbf{A}^{(1)}(R)$

The calculation of matrix $\mathbf{A}^{(1/2)}(R)$ introduced by (3.2) involves more difficulties. In the present subsection, we shall compute, instead of $\mathbf{A}^{(1/2)}(R)$, the matrix $\mathbf{A}^{(1)}(R)$ defined by

$$\mathbf{A}^{(1)}(R)\equiv\mathbf{D}^{-1}\tilde{\mathbf{U}}(R)\mathbf{D}=\sum_P a_P{}^{(1)}(R)\tilde{\mathbf{U}}(P). \quad (3.18)$$

In subsection (b), the matrix $\mathbf{A}^{(1/2)}(R)$ will be calculated by using the expansion coefficients $a^{(1)}(R)$ obtained here.

We shall proceed with the calculation of $\mathbf{A}^{(1)}(R)$ by observing from (3.18) that $\mathbf{A}^{(1)}(R)$ can be expanded in terms of matrices $\mathbf{F}$ and consequently $a_P{}^{(1)}(R)$ in terms of f_P. To facilitate the expansion of $\mathbf{A}^{(1)}(R)$ in terms of $\mathbf{F}$, we shall consider the Hermitian conjugate of the expression (3.18), that is,

$$\mathbf{A}^{(1)}(R^{-1})^\dagger=\mathbf{D}\tilde{\mathbf{U}}(R)\mathbf{D}^{-1}=\sum_P a_P{}^{-1(1)}(R^{-1})^*\tilde{\mathbf{U}}(P). \quad (3.19)$$

Let us first consider a case where R is a simple interchange (h_1h_2). Use of the expansion (2.21) of $\mathbf{D}$ in $\mathbf{D}\tilde{\mathbf{U}}(R)\mathbf{D}^{-1}$ decomposes the matrix $\mathbf{A}^{(1)}(h_1h_2)^\dagger$ into a sum of matrices of the type

$$\langle k_1|h_1\rangle\mathbf{S}[k_1h_2|h_1h_2]\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}. \quad (3.20)$$

When the method used in obtaining the expansion (3.16) is applied in $\mathbf{S}[k_1h_2|h_1h_2]$, the above expression is rewritten as

---

$$\begin{aligned}
&\langle h_1|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1)\mathbf{S}[h_1h_2k_1|h_1h_2k_1]\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}\\
&\quad+\sum_{k_2}{}''\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1k_2)\mathbf{S}[h_1h_2k_1k_2|h_1h_2k_1k_2]\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}\\
&\quad+\sum_{k_3}{}'''\sum_{k_2}{}''\langle h_1|k_3\rangle\langle k_3|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1k_2k_3)\mathbf{S}[h_1h_2k_1k_2k_3|h_1h_2k_1k_2k_3]\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}\\
&\quad+\cdots\\
&=\langle h_1|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1h_2)\mathbf{F}[h_1h_2k_1|h_1h_2k_1]\\
&\quad+\sum_{k_2}{}'''\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1k_2h_2)\mathbf{F}[h_1h_2k_1k_2|h_1h_2k_1k_2]\\
&\quad+\sum_{k_3}{}''''\sum_{k_2}{}'''\langle h_1|k_3\rangle\langle k_3|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\tilde{\mathbf{U}}(h_1k_1k_2k_3h_2)\mathbf{F}[h_1h_2k_1k_2k_3|h_1h_2k_1k_2k_3]\\
&\quad+\cdots. \qquad (3.21)
\end{aligned}$$

Similarly $\mathbf{D}$ in the expression $\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}\mathbf{D}^{-1}$ is expanded, where a term corresponding to (3.20) yields a series similar to (3.21), but $\tilde{\mathbf{U}}(h_1h_2k_1)$, $\tilde{\mathbf{U}}(h_1h_2k_1k_2)$, etc., will appear instead of $\tilde{\mathbf{U}}(h_1k_1h_2)$, $\tilde{\mathbf{U}}(h_1k_1k_2h_2)$, etc., in (3.21). The matrix

$$\langle k_1|h_2\rangle\mathbf{S}[h_1k_1|h_1h_2]\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}$$

will also produce an expansion obtained from (3.21) by interchanging h_1 and h_2.

By calculating all matrices appearing in the expansion of $\mathbf{D}$ and by taking the difference between $\mathbf{D}\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and $\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}\mathbf{D}^{-1}=\tilde{\mathbf{U}}(h_1h_2)$, it will finally be found that

$$\begin{aligned}\mathbf{A}^{(1)}(h_1h_2)^\dagger&=\mathbf{D}\tilde{\mathbf{U}}(h_1h_2)\mathbf{D}^{-1}\\&=\tilde{\mathbf{U}}(h_1h_2)+\sum_{k_1}{}'(\langle h_1|k_1\rangle\langle k_1|h_1\rangle-\langle h_2|k_1\rangle\langle k_1|h_2\rangle)[\tilde{\mathbf{U}}(h_1k_1h_2)-\tilde{\mathbf{U}}(h_2k_1h_1)]\mathbf{F}[h_1h_2k_1|h_1h_2k_1]\\&\quad+\sum_{k_1}{}''(\langle h_1|k_1\rangle\langle k_1|h_2\rangle\langle h_2|h_1\rangle-\langle h_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|h_2\rangle)[\tilde{\mathbf{U}}(h_2k_1)-\tilde{\mathbf{U}}(h_1k_1)]\mathbf{F}[h_1h_2k_1|h_1h_2k_1]\\&\quad+\sum_{k_2}{}'''\sum_{k_1}{}''(\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle-\langle h_2|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_2\rangle)\\&\qquad\times[\tilde{\mathbf{U}}(h_1k_1k_2h_2)-\tilde{\mathbf{U}}(h_2k_1k_2h_1)]\mathbf{F}[h_1h_2k_1k_2|h_1h_2k_1k_2]\\&\quad+\cdots\\&\quad+\sum_{k_2}{}'''\sum_{k_1}{}''(\langle h_1|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_2\rangle\langle h_2|h_1\rangle-\langle h_2|k_2\rangle\langle k_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|h_2\rangle)\\&\qquad\times[\tilde{\mathbf{U}}(h_2k_1k_2)-\tilde{\mathbf{U}}(h_1k_1k_2)]\mathbf{F}[h_1h_2k_1k_2|h_1h_2k_1k_2]\\&\quad+\cdots\\&\quad+\tfrac{1}{2}\sum_{k_2}{}'''\sum_{k_1}{}''(\langle h_1|k_1\rangle\langle k_1|h_1\rangle\langle h_2|k_2\rangle\langle k_2|h_2\rangle-\langle h_2|k_1\rangle\langle k_1|h_2\rangle\langle h_1|k_2\rangle\langle k_2|h_1\rangle)\\&\qquad\times[\tilde{\mathbf{U}}(h_1k_1h_2k_2)-\tilde{\mathbf{U}}(h_2k_1h_1k_2)]\mathbf{F}[h_1h_2k_1k_2|h_1h_2k_1k_2]\\&\quad+\cdots\\&\quad+\tfrac{1}{2}\sum_{k_2}{}'''\sum_{k_1}{}''(\langle h_1|k_1\rangle\langle k_1|h_2\rangle\langle h_2|k_2\rangle\langle k_2|h_1\rangle-\langle h_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|k_2\rangle\langle k_2|h_2\rangle)\\&\qquad\times[\tilde{\mathbf{U}}((h_1k_2)(h_2k_1))-\tilde{\mathbf{U}}((h_1k_1)(h_2k_2))]\mathbf{F}[h_1h_2k_1k_2|h_1h_2k_1k_2]\qquad(3.22)\\&\quad+\cdots.\end{aligned}$$

By using the expansion technique of the $\mathbf{F}$ matrices developed in the previous subsection (1), therefore, the matrix $\mathbf{A}^{(1)}(h_1h_2)$ can be calculated. The inequality (A7) ensures the convergence even if higher order terms in n of $\mathbf{F}[h(n)]$ are neglected.

Let us insert (3.1) and (3.17) into (3.22) and write the series in terms of loops S_P and $\mathbf{U}(P)$ as follows:

$$\begin{aligned}&\mathbf{A}^{(1)}(h_1h_2)^\dagger\\&\quad=\tilde{\mathbf{U}}(h_1h_2)+\sum_P\sum_t\sum_{P_0P_1\cdots P_n}^{P}(-1)^tC[(h_1h_2);P_0]\\&\qquad\times S_{P_1}S_{P_2}\cdots S_{P_t}\tilde{\mathbf{U}}(P),\qquad(3.23)\end{aligned}$$

where the summation $\sum_{P_0P_1\cdots P_t}{}^P$ takes all possible $P_0P_1\cdots P_t$ which satisfy the relation

$$P=P_tP_{t-1}\cdots P_2P_1(h_1h_2)P_0,\qquad(3.24)$$

and $C[(h_1h_2);P_0]$ represents a quantity in parentheses in (3.22). For instance,

$$\begin{aligned}&C[(h_1h_2);(h_1k_1)]=\langle h_1|k_1\rangle\langle k_1|h_1\rangle-\langle h_2|k_1\rangle\langle k_1|h_2\rangle,\\&C[(h_1h_2);(h_1h_2k_1)]=\langle h_1|k_1\rangle\langle k_1|h_2\rangle\langle h_2|h_1\rangle\\&\qquad-\langle h_2|k_1\rangle\langle k_1|h_1\rangle\langle h_1|h_2\rangle.\qquad(3.25)\end{aligned}$$

Comparison of (3.23) with (3.19) yields that

$$\begin{aligned}&a_{P^{-1}}{}^{(1)}(h_1h_2)^*\\&\quad=\delta_{(h_1h_2),P^{-1}}+\sum_t\sum_{P_0P_1\cdots Pt}^{P}(-1)^tC[(h_1h_2);P_0]\\&\qquad\times S_{P_1}S_{P_2}\cdots S_{P_t}.\qquad(3.26)\end{aligned}$$

By noting the relation $S_P{}^*=S_{P^{-1}}$, the above equation can be rewritten as

$$\begin{aligned}&a_P{}^{(1)}(h_1h_2)=\delta_{(h_1h_2),P}\\&\qquad+\sum_t\sum_{P_0P_1\cdots P_n}^{P}{}'(-1)^tC[(h_1h_2);P_0]^*\\&\qquad\times S_{P_1}{}^{-1}S_{P_2}{}^{-1}\cdots S_{P_t}{}^{-1},\qquad(3.27)\end{aligned}$$

where the summation $\sum_{P_0P_1\cdots P_t}{}'$ takes all possible $P_0, P_1\cdots P_t$ under the condition

$$P=P_0{}^{-1}(h_1h_2)P_1{}^{-1}P_2{}^{-1}\cdots P_t{}^{-1}.\qquad(3.28)$$

The equation (3.27), however, is less convenient than (3.26). To evaluate the electron permutation P as given by (3.28), we have to start from the last loop S_{P_t} and to count the cyclic permutations in reverse order. On the other hand, computation of $a_{P^{-1}}{}^{(1)}(h_1h_2)^*$ by (3.26) and (3.24) is exactly the same as the expansion (3.17) of $f_P[h(n)]$ and the electron permutation P can be obtained by multiplying cyclic permutations following the same order in which the loops appear. It is easily found that, after completing a diagram of $\mathbf{A}_{P^{-1}}{}^{(1)}(h_1h_2)^*$ by (3.26), a corresponding diagram for $\mathbf{A}_P{}^{(1)}(h_1h_2)$ can be obtained by reversing the direction of arrows in all solid lines as well as broken lines.

The calculation of $\mathbf{A}^{(1)}(R)$ just described can be extended to any permutation R, and it will be found that

$$\begin{aligned}a_{P^{-1}}{}^{(1)}(R^{-1})^*&=\delta_{R,P}+\sum_t\sum_{P_0P_1\cdots Pt}^{P}(-1)^tC[R;P_0]\\&\qquad\times S_{P_1}S_{P_2}\cdots S_{P_t}.\qquad(3.29)\end{aligned}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Diagrams for $a_P^{(1)}(R^{-1})$ can be constructed from those obtained by (3.29), by simply reversing directions of arrows in all solid lines and broken lines.

The only problem here is to develop $\mathbf{A}^{(1)}(R^{-1})^\dagger$ in terms of $\mathbf{F}$ and get an equation similar to (3.22), which, in return, gives explicit expressions of $C[R; P_0]$. This can also be carried out automatically by drawing diagrams of $C[R; P_0]$ as follows: Let R be a cyclic permutation $R=(h_1h_2\cdots h_s)$ and consider all possible permutations P_0 which start from a point on the loop $(h_1h_2\cdots h_s)$. The permutations P_0 may consist of many unconnected loops $P_{01}, P_{02}, \cdots, P_{0r}$, but each of the loops must have at least one common point with the original loop $(h_1h_2\cdots h_s)$. The desired coefficients $C[R; P_0]$ will be given by

$$C[R; P_0]=S_{P_{01}}S_{P_{02}}\cdots S_{P_{0r}} -S_{R^{-1}P_{01}R}S_{R^{-1}P_{02}R}\cdots S_{R^{-1}P_{0r}R}, \quad (3.30)$$

where $P_0=P_{0r}P_{0r-1}\cdots P_{02}P_{01}$.

In case R is a single interchange, we had the relation

$$C[R; P_0]=-C[R; R^{-1}P_0R], \quad (3.31)$$

which leads to factorizations of the type

$$C[R; P_0][\tilde{\mathbf{U}}(R^{-1}P_0)-\tilde{\mathbf{U}}(P_0R^{-1})], \quad (3.32)$$

as has been shown in (3.22). When R is a permutation involving more than two electrons, the relation (3.31) is no longer valid and each matrix $\tilde{\mathbf{U}}(R^{-1}P_0)$ may carry a distinct coefficient $C[R; P_0]$. Therefore $\mathbf{A}^{(1)}(R^{-1})^\dagger$ will not be written in the form (3.32).

Use of the inequality (A7) in (3.22) leads to the conclusion that the sum

$$L_b^{(2)}=\sum_P |a_P^{(1)}(h_1h_2)-\delta_{(h_1h_2),P}| \quad (3.33)$$

is also bounded and small as compared with unity as long as Δ is small. However, the quantity

$$L_b^{(s)}=\sum_P |a_P^{(1)}(R_s^{-1})-\delta_{R_s^{-1},P}| \quad (3.34)$$

will increase in proportion to the number s of electrons involved in the permutations R_s. This will be seen if we calculate some of the dominant terms in $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$. For instance, let R_s be a loop $(h_1h_2\cdots h_s)$ of equally spaced s electrons in a crystal with translational symmetry. The matrix $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$ will include s terms of the form

$$C[R_s; (h_ik_i)]\mathbf{U}(R_s^{-1}(h_ik_i)), \quad \text{for} \quad i=1, 2, \cdots, n,$$

whose s coefficients

$$C[R_s; (h_ik_i)]=\langle h_i|k_i\rangle\langle k_i|h_i\rangle-\langle h_{i+1}|k_i\rangle\langle k_i|h_{i+1}\rangle,$$

are equal to each other. Terms of the form

$$C[R_s; (h_ik_i)(h_jk_j)]\mathbf{U}(R_s^{-1}(h_ik_i)(h_jk_j))$$

will appear $\frac{1}{2}s(s-1)$ times in $A^{(1)}(R_s^{-1})^\dagger$, since all possible relative positions of loops (h_ik_i) and (h_jk_j) should be considered here.

This divergence of $L_b^{(s)}$ will lead to difficulties in proving the convergence of the expansion of $\mathbf{A}^{(1/2)}(R)$ as is seen in the following subsection (b). The origin of the divergence is due to the fact that diagrams in $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$ are not connected with solid lines (overlaps). It is expected that, if the quantity $S_{R_s}\mathbf{A}^{(1)}(R_s)^\dagger$ is considered instead of $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$, the sum of the absolute values of the expansion coefficients $S_{R_s}a_P^{(1)}(R_s^{-1})$

$$M_b^{(s)}=\sum_P |S_{R_s}[a_P^{(1)}(R_s^{-1})-\delta_{R_s^{-1}P}]| \leqslant |S_{R_s}| L_b^{(s)} \quad (3.35)$$

will converge and be very small. In fact, there is no essential difference between $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$ and $\mathbf{F}[h(n)]$, and their expansion coefficients are calculated similarly by (3.17) and (3.26). Therefore the argument used in Appendix for bounding $F[h(n)]$ will be applied to $\mathbf{A}^{(1)}(R_s^{-1})^\dagger$ too.

(b) *Asymptotic Expansion of* $\mathbf{A}^{(1/2)}(B)$.

The expansion coefficients $a_P^{(1)}(R)$ obtained in the preceding subsection (a) are related to the desired ones $a_P^{(1/2)}(R)$ by

$$a_P^{(1)}(R)=\sum_Q a_Q^{(1/2)}(R)a_P^{(1/2)}(Q). \quad (3.36)$$

This relation is proved as follows. The matrix $\mathbf{A}^{(1)}(R)$ defined by (3.18) can also be expanded as

$$\begin{aligned}\mathbf{A}^{(1)}(R)&=\mathbf{D}^{-1/2}\mathbf{D}^{-1/2}\tilde{\mathbf{U}}(R)\mathbf{D}^{1/2}\mathbf{D}^{1/2}\\ &=\sum_Q a_Q^{(1/2)}(R)\mathbf{D}^{-1/2}\tilde{\mathbf{U}}(Q)\mathbf{D}^{1/2}\\ &=\sum_P\sum_Q a_Q^{(1/2)}(R)a_P^{(1/2)}(Q)\tilde{\mathbf{U}}(P). \quad (3.37)\end{aligned}$$

Comparison of (3.37) with (3.18) yields (3.36).

Let $a_P^{(1/2)}(R)$ and $a_P^{(1)}(R)$ be RP elements of matrices $\mathbf{a}^{(1/2)}$ and $\mathbf{a}^{(1)}$, respectively, and rewrite Eq. (3.36) in terms of the matrix representation as follows:

$$\mathbf{a}^{(1)}=\mathbf{a}^{(1/2)}\mathbf{a}^{(1/2)}. \quad (3.38)$$

Use of the orthogonality relation of the irreducible representation matrices $\tilde{\mathbf{U}}(P)$ in (3.2) leads to[21,22]

$$a_P^{(1/2)}(R)=\sum_{\rho km}\frac{f_\rho}{N!}(D^{-1/2})_{kk}^{(\rho)}\tilde{U}_{km}^{(\rho)}(R) \times(D^{1/2})_{mm}^{(\rho)}\tilde{U}_{km}^{(\rho)}(P)^*, \quad (3.39)$$

[21] As we shall show in (4.1) and (4.2), it is always possible to consider a representation where $\mathbf{D}$ is diagonal.

[22] The $N!$ matrices $\tilde{\mathbf{U}}(P)$ introduced by (2.8) are representation matrices of the permutation group. By choosing a suitable set of N-spin functions Θ_i as basis, $\tilde{\mathbf{U}}(P)$'s can be made irreducible. Let $\tilde{\mathbf{U}}^{(\rho)}(P)$ be the matrix in the irreducible representation ρ. The km

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where f_ρ is the dimension of the subspace (ρ). Since $(D^{-1/2})_{kk}{}^{(\rho)}$ and $(D^{1/2})_{mm}{}^{(\rho)}$ are real, we find the relation

$$a_P{}^{(1/2)}(R)=a_R{}^{(1/2)}(P)^*, \tag{3.40}$$

which shows that the matrices $\mathbf{a}^{(1/2)}$ and $\mathbf{a}^{(1)}$ are both Hermitian. Hence, there exists a unitary matrix $\mathbf{V}$ with which $\mathbf{a}^{(1/2)}$ and $\mathbf{a}^{(1)}$ are simultaneously diagonalized such that

$$\mathbf{V}^\dagger\mathbf{a}^{(1)}\mathbf{V}=\mathbf{1}+\boldsymbol{\eta}, \quad \text{diagonal}, \tag{3.41}$$

$$\mathbf{V}^\dagger\mathbf{a}^{(1/2)}\mathbf{V}=(\mathbf{1}+\boldsymbol{\eta})^{1/2}, \quad \text{diagonal}. \tag{3.42}$$

Eigenvalues $1+\eta_K$ and $(1+\eta_K)^{1/2}$ of matrices $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(1/2)}$ are all positive, and there is no case where $(\mathbf{1}+\boldsymbol{\eta})^{1/2}$ on the right of (3.42) carries a minus sign. This will be proved as follows. Let us introduce matrix $\mathbf{A}^{(1/4)}(R)$ and the expansion coefficients $a_P{}^{(1/4)}(R)$ by

$$\mathbf{A}^{(1/4)}(R)\equiv\mathbf{D}^{-1/4}\tilde{\mathbf{U}}(R)\mathbf{D}^{1/4}=\sum_P a_P{}^{(1/4)}(R)\tilde{\mathbf{U}}(P). \tag{3.43}$$

As before, a Hermitian matrix $\mathbf{a}^{(1/4)}$ will be constructed from the expansion coefficients $a_P{}^{(1/4)}(R)$, and the following relation between $\mathbf{a}^{(1/4)}$ and $\mathbf{a}^{(1/2)}$ will be found

$$\mathbf{a}^{(1/2)}=\mathbf{a}^{(1/4)}\mathbf{a}^{(1/4)}. \tag{3.44}$$

This relation shows that $\mathbf{a}^{(1/4)}$ is also brought into a diagonal form by the unitary matrix $\mathbf{V}$ used in (3.41) and (3.42), and that the eigenvalues are square roots of eigenvalues of $\mathbf{a}^{(1/2)}$. This implies that, if some eigenvalues of $\mathbf{a}^{(1/2)}$ carried minus signs, the corresponding eigenvalues of $\mathbf{a}^{(1/4)}$ would have to be imaginary. This is contradictory to the theorem that the eigenvalues of a Hermitian matrix are real. Therefore, eigenvalues $1+\eta_K$ and $(1+\eta_K)^{1/2}$ in (3.41) and (3.42) must be all positive.

Since off-diagonal elements $a_P{}^{(1)}(R)$ are small, the relations (3.41) and (3.42) suggest that the perturbation technique can be applied to calculate the eigenvalues $1+\eta_K$ and the unitary matrix $\mathbf{V}$ of $\mathbf{a}^{(1)}$. The desired matrix $\mathbf{a}^{(1/2)}$ can then be obtained from

$$\mathbf{a}^{(1/2)}=\mathbf{V}(\mathbf{1}+\boldsymbol{\eta})^{1/2}\mathbf{V}^\dagger. \tag{3.45}$$

It turned out that this approach is not practical since the matrix $\mathbf{a}^{(1)}$ has an infinite order of degeneracy. For instance, diagonal elements $a_{(h_ik_i)}{}^{(1)}(h_ik_i)$, for $i=1, 2, \cdots, N$, are all equal to each other, if we consider a crystal with translational symmetry.

To eliminate this difficulty, it is more convenient to use an asymptotic expansion. Let us expand $(\mathbf{1}+\boldsymbol{\eta})^{1/2}$ in (3.45) in a power series of $\boldsymbol{\eta}$. The matrix $\mathbf{a}^{(1/2)}$ is then written formally as

$$\begin{aligned}\mathbf{a}^{(1/2)}&=\mathbf{V}(\mathbf{1}+\tfrac{1}{2}\boldsymbol{\eta}-\tfrac{1}{2}\times\tfrac{1}{4}\boldsymbol{\eta}^2+\tfrac{1}{2}\times\tfrac{1}{4}\times(3/6)\boldsymbol{\eta}^3-\cdots)\mathbf{V}^\dagger\\&=\mathbf{1}+\tfrac{1}{2}\mathbf{b}-\tfrac{1}{2}\times\tfrac{1}{4}\mathbf{b}^2+\tfrac{1}{2}\times\tfrac{1}{4}\times(3/6)\mathbf{b}^3-\cdots,\end{aligned} \tag{3.46}$$

where

$$\mathbf{b}\equiv\mathbf{V}\boldsymbol{\eta}\mathbf{V}^\dagger=\mathbf{a}^{(1)}-\mathbf{1}. \tag{3.47}$$

These equations (3.46) and (3.47) are written more explicitly as

$$a_P{}^{(1/2)}(R)=\delta_{RP}+\tfrac{1}{2}b_{RP}-\tfrac{1}{2}\times\tfrac{1}{4}\sum_Q b_{RQ}b_{QP}+\cdots, \tag{3.48}$$

and

$$b_{RP}=a_P{}^{(1)}(R)-\delta_{RP}. \tag{3.49}$$

Since b_{RP} is obtained by (3.29) and (3.49), the expansion coefficients $a_P{}^{(1/2)}(R)$ can be calculated by (3.48) provided the series converges.

In practice, there will not be any difficulty in using the expansion (3.48) for the calculation of $a_P{}^{(1/2)}(R)$. This will be explained as follows: Since it is not possible to calculate b_{RP} for all R and P exactly, we need to truncate and approximate the matrix b by neglecting permutations involving $s+1$ electrons or more. Let b^0 be the truncated matrix. The RP element $b_{RP}{}^0$ is then given by an approximate b_{RP} when R and P are both permutations of s electrons or less, while $b_{RP}{}^0=0$ if either R or P involves $s+1$ electrons or more. Here s is chosen so that $L_b{}^{(s)}$ as defined by (3.34) is less than one. This implies that

$$\sum_P|b_{RP}{}^0|\leqslant\epsilon<1, \quad \text{for all } R. \tag{3.50}$$

When all of b_{RP} in (3.48) are replaced by $b_{RP}{}^0$, the series (3.48) converges and the quantities $a_P{}^{(1/2)\,(0)}(R)$ thus calculated satisfy the following inequality:

$$\sum_P|a_P{}^{(1/2)\,(0)}(R)|\leqslant 2-(1-\epsilon)^{1/2}<(1-\epsilon)^{-1/2}. \tag{3.51}$$

The relation (3.51) will be proved in exactly the same way as (A6) has been obtained.

When the number s increases, however, the value of $L_b{}^{(s)}$ increases indefinitely, and the convergence of (3.48) will not be proved easily by an exact language of algebra.[11] Nevertheless, it is expected that the series (3.48), multiplied by S_R

$$S_Ra_P{}^{(1/2)}(R)=S_R\delta_{RP}+\tfrac{1}{2}S_Rb_{RP}-\tfrac{1}{2}\times\tfrac{1}{4}\sum_Q S_Rb_{RQ}b_{QP}+\cdots \tag{3.52}$$

converges, for the following reasons. Quantities such as S_Rb_{RP} or $\sum_Q S_Rb_{RQ}$ will be finite and small as we have discussed in (3.35). Higher terms $\sum_Q S_Rb_{RQ}b_{QP}$, $\sum_{Q_2}\sum_{Q_1}S_Rb_{RQ_1}b_{Q_1Q_2}b_{Q_2P}$, etc., consist of connected dia-

---

element of matrix $\mathbf{A}^{(1/2)}(R)$ in this representation is then written as

$$[A^{(1/2)}(R)]_{km}{}^{(\rho)}\equiv(D^{-1/2})_{kk}{}^{(\rho)}\tilde{\mathbf{U}}_{km}{}^{(\rho)}(R)(D)_{mm}{}^{(\rho)}=\sum_Q a_Q{}^{(1/2)}(R)\tilde{\mathbf{U}}_{km}{}^{(\rho)}(Q). \tag{1}$$

Because of the orthogonality relation of the irreducible representation matrices

$$\sum_{\rho km}(f_\rho/N!)\tilde{\mathbf{U}}_{km}{}^{(\rho)}(Q)\tilde{\mathbf{U}}_{km}{}^{(\rho)}(P)^*=\delta_{QP}, \tag{2}$$

we will find (3.39) when multiplying (1) by $\tilde{\mathbf{U}}_{km}{}^{(\rho)}(P)^*$ and summing over ρ, k and m.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

grams, which build up over $S_R b_{RQ}$. This implies the convergence of the series provided the same diagrams do not accumulate. When a diagram appears repeatedly, we can count the number of times the diagram appears as we did for the $\mathbf{F}$ matrix in Sec. III (1b), and it will be finally found that they cancel each other in the course of the summation in (3.52), so the net contribution is a small finite number.

The calculation of $\sum_Q S_R b_{RQ} b_{QP}$ is, in fact, very similar to that of $S_{R_n}\mathbf{F}[h(n)]$ described by (3.15) or (3.9), where b_{RQ} corresponds to $\Gamma[k(n)]$ and b_{QP} to $\mathbf{F}_{h(n)}[k(m)]$. $S_{R_n}F[h(n)]$ may contain very large terms $\mathbf{F}_{h(n)}[k(m)]$ when m is large, but those are connected by $S_{R_n}\Gamma[k(m)]$. Large terms $b_{Q_m P}$ for large m are also connected by $S_R b_{RQ_m}$. To evaluate higher terms $\sum_{Q_2}\sum_{Q_1} S_R b_{RQ_1} b_{Q_1 Q_2} b_{Q_2 P}$, etc., corresponds to expanding $\mathbf{F}_{h(n)}[k(m)]$ in terms of $\mathbf{F}_{h(n)k(m)}[l(s)]$, etc. As we have discussed in the Appendix, the value of $S_{R_n}\mathbf{F}[h(n)]$ vanishes when n becomes large.

The higher order contribution, for which $L_b^{(s)} > 1$, can also be calculated by combining a perturbation

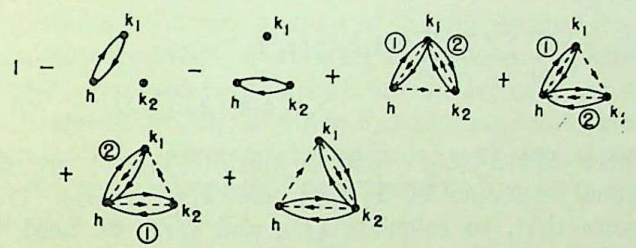


FIG. 7. The expansion of $\mathbf{F}[h|h]$.

technique with the asymptotic expansion described here. We shall not describe the details but we will again reach the same conclusion that $S_R\mathbf{A}^{(1/2)}(R)$ is expanded in terms of connected diagrams.

(c) *Role of* $\mathbf{A}^{(1/2)}(R)$

We have described the matrix $\mathbf{A}^{(1/2)}(R)$ as being very close to one and merely a correction factor of the $\mathbf{F}$ matrix. Here we shall illustrate, by an example, an essential role of $\mathbf{A}^{(1/2)}$ in our calculation.

Let us consider a Hermitian matrix

$$\mathbf{S}_N = \mathbf{D}^{-1/2}\mathbf{S}[h|h]\mathbf{D}^{-1/2},$$

which, according to (3.3), is expanded as

$$\mathbf{S}_N = \mathbf{D}^{-1/2}\mathbf{S}[h|h]\mathbf{D}^{-1/2} = \sum_P \sum_Q f_Q[h|h] a_P^{(1/2)}(Q)\tilde{\mathbf{U}}(P). \quad (3.53)$$

By using the prescription (i)–(iv) in Sec. III (1b) for drawing diagrams of matrix $\mathbf{F}[h|h]$, we find that $\mathbf{F}[h|h]$ contains among others the diagrams illustrated in Fig. 7.

If we set $\mathbf{A}^{(1/2)}(R) = 1$ and hence $a_P^{(1/2)}(R) = \delta_{RP}$, the matrix $\mathbf{S}_N$ becomes equivalent to $\mathbf{F}[h|h]$ described in Fig. 7, which is not Hermitian. For example, Hermitian

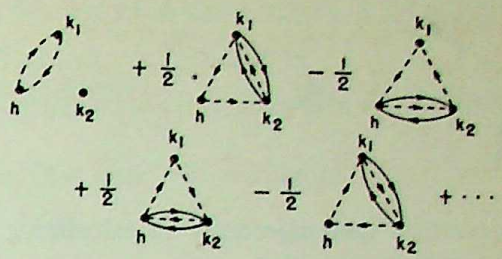


FIG. 8. The expansion of $\mathbf{A}^{(1/2)}(hk_1)$.

conjugates of the fourth and seventh diagrams become equivalent to, respectively, the second and fourth diagrams in Fig. 9. However, those diagrams cannot appear in $\mathbf{F}[h|h]$, since, to draw either of them, we have to start from loop (k_1k_2). All diagrams in $\mathbf{F}[h|h]$, however, should start from h.

On the other hand, matrix $\mathbf{A}^{(1/2)}(hk_1)$ includes the diagrams in Fig. 8. Diagrams for $\mathbf{A}^{(1/2)}(hk_2)$ are similar to those in Fig. 8 and obtained from them by interchanging k_1 and k_2. If we multiply $\mathbf{A}^{(1/2)}(hk_1)$ by the second term in Fig. 7 and $\mathbf{A}^{(1/2)}(hk_2)$ by the third, we obtain the expression for $\mathbf{S}_N$, which is similar to the series in Fig. 7, but the fourth and seventh terms are replaced by those in Fig. 9, and matrix $\mathbf{S}_N$ becomes Hermitian.

The present example clearly shows that the matrix $\mathbf{A}^{(1/2)}(R)$ is essential for retaining the Hermitian property of the energy matrix.

IV. DISCUSSION

(1) Evaluation of the Expectation Values

According to the description in Sec. III, any matrix containing an operator $f(\mathbf{r}_h)$ or $g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})$ can now be evaluated by drawing diagrams. Let us denote an integral of the form $\langle k|f(\mathbf{r}_h)|h\rangle$ by a heavy line starting from h and ending at k plus a wavy line indicating the operator $f(\mathbf{r}_h)$ as shown in Fig. 10, where the second one corresponds to $\langle h|f(\mathbf{r}_h)|h\rangle$. The integral

$$\langle k_1k_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$$

is also written by diagrams as in Fig. 11. Here the second one represents a Coulomb interaction, the third one an exchange interaction.

From (3.16) and (3.17) together with the discussion at the end of Sec. III (2b), it is clear that any expectation value $\langle\mathbf{f}\rangle$ or $\langle\mathbf{g}\rangle$ given by (2.26) or (2.27) can be written as a sum of connected diagrams. This provides another indication that the nonorthogonality difficulties do not appear and that the Heitler–London

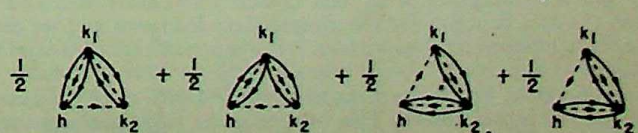


FIG. 9. Terms appearing in the expansion of $\mathbf{S}_N = \mathbf{D}^{-1/2}\mathbf{S}[h|h]\mathbf{D}^{-1/2}$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

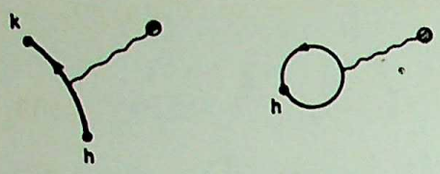


FIG. 10. Diagrams for $\langle k|f(\mathbf{r}_h)|h\rangle$ and $\langle h|f(\mathbf{r}_h)|h\rangle$.

method can be used rigorously in calculating energy spectra or other quantities.

The calculation can be carried out as follows:

(i) First, we shall write an interaction term $\langle k|f(\mathbf{r})|h\rangle$ or $\langle k_1k_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$ by using notations introduced in Figs. 10 and 11.

(ii) The matrix $\mathbf{D}^{-1/2}\mathbf{S}[k|h]\mathbf{D}^{-1/2}$ or

$$\mathbf{D}^{-1/2}S[k_1k_2|h_1h_2]\mathbf{D}^{-1/2}$$

should follow the diagram just obtained in (i). These matrices are expanded in terms of $f_P[h(n)]$ and $a_P{}^{(1/2)}(R)$.

(iii) Diagrams for $f_P[h(n)]$ are all constructed according to the prescriptions (i)–(iv) in Sec. III (1b).

(iv) Before computing $a_P{}^{(1/2)}(R)$, we need to draw diagrams for $a_P{}^{(1)}(R)$. This will be accomplished according to the instructions appearing between (3.29) and (3.30).

(v) The desired coefficients $a_P{}^{(1/2)}(R)$ are computed by the asymptotic expansion (3.48) and (3.49).

(vi) Finally, the matrix $\mathbf{D}^{-1/2}\mathbf{S}[k|h]\mathbf{D}^{-1/2}$ or

$$\mathbf{D}^{-1/2}\mathbf{S}[k_1k_2|h_1h_2]\mathbf{D}^{-1/2}$$

is calculated by (3.3).

(vii) Repeat the process for all k. Thus, the calculation of $\langle\mathbf{f}\rangle$ or $\langle\mathbf{g}\rangle$, including the inverse of the overlap, is now completed.

(2) Calculation of the Ground-State Energy

So far, our aim was to calculate spin-wave spectra of the Heitler–London space. The wave function $\mathbf{\Psi}$ defined by (2.3), therefore, was a vector in the complete space G of spin waves, and the energy matrix had the dimensions of the space G. This matrix representation has introduced the matrices $\mathbf{A}^{(1/2)}(R)$, which are difficult to calculate.

If, on the other hand, we want to compute the ground state of the Heitler–London method by assuming a definite spin function Θ^0 for the system,[23] we can bypass the difficult calculation of $\mathbf{A}^{(1/2)}(R)$ and we need to compute the $\mathbf{F}$ matrices in one dimension, only.

The 2^N wave functions Ψ_i defined by (2.3) can be regarded as orthogonal and the overlap matrix $\mathbf{D}$ as diagonal. If $\mathbf{D}$ is not diagonal, there exists a unitary matrix $\mathbf{V}$ with which $\mathbf{D}$ can be brought into diagonal form such that

$$\mathbf{D}'=\mathbf{V}^\dagger\mathbf{D}\mathbf{V}=\int(\mathbf{\Psi}\mathbf{V})^\dagger(\mathbf{\Psi}\mathbf{V})d\tau. \qquad (4.1)$$

This implies that, instead of $\mathbf{\Psi}$, we should use the orthogonal wave functions $\mathbf{\Psi}\mathbf{V}$ defined by

$$\mathbf{\Psi}\mathbf{V}=(N!)^{-1/2}\sum_P(-1)^PP^\tau X(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N)\times P^\sigma\mathbf{\Theta}(\sigma_1,\sigma_2,\cdots,\sigma_N)\mathbf{V}, \qquad (4.2)$$

and, instead of $\mathbf{\Theta}$, spin functions $\mathbf{\Theta}\mathbf{V}$ should be used as basis.

Use of the diagonal matrix $\mathbf{D}$ in (2.26) and (2.27) leads to

$$\langle f\rangle_{ii}=\sum_h\langle h|f(\mathbf{r}_h)|h\rangle S[h|h]_{ii}D_{ii}{}^{-1}+\sum_h\sum_k{}'\langle k|f(\mathbf{r}_h)|h\rangle S[k|h]_{ii}D_{ii}{}^{-1}, \qquad (4.3)$$

and

$$\langle g\rangle_{ii}=\tfrac{1}{2}\sum_{h_1}\sum_{h_2}{}'\sum_{k_1}\sum_{k_2}{}'\langle k_1k_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle\times S[k_1k_2|h_1h_2]_{ii}D_{ii}{}^{-1}, \qquad (4.4)$$

where $S[k|h]_{ii}D_{ii}{}^{-1}$ and $S[k_1k_2|h_1h_2]_{ii}D_{ii}{}^{-1}$ are diagonal elements of $\mathbf{F}[k|h]$ and $\mathbf{F}[k_1k_2|h_1h_2]$. This shows that, to calculate $\langle f\rangle_{ii}$ and $\langle g\rangle_{ii}$, we need to compute the $\mathbf{F}$ matrices only.

The overlap matrices $\mathbf{D}$ and $\mathbf{S}$ introduced by (2.7) and (2.18) can be regarded as if they are one dimensional. Hence matrices $\mathbf{F}$ and $\mathbf{X}$ introduced in Sec. III are all one dimensional and $\mathbf{U}(P)$ is a number defined by (2.6) where $i=j$. The description of drawing diagrams of the $\mathbf{F}$ matrices in Sec. III (1b) is now applied to this "one dimensional" case, and the calculation will be simplified considerably.

(3) Intra-Atomic Terms

As we have remarked in the end of Sec. I, our method can be applied to a case where each of the atoms carries more than one electron. We need to make only two short remarks for this case.

Since atomic orbitals centered at a nucleus are orthogonal to each other, overlap integrals between them vanish and we need to draw lines between those belonging to different atoms. It is convenient, therefore,

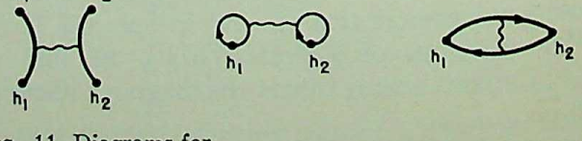


FIG. 11. Diagrams for $\langle k_1k_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$, $\langle h_1h_2|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$ and $\langle h_2h_1|g(\mathbf{r}_{h_1}\mathbf{r}_{h_2})|h_1h_2\rangle$.

[23] For a ferromagnet or an ionic crystal with closed-shell structure, the spin function for the ground state is known and we can use this simplified method without introducing further approximations. For any other type of ground state, we have to assume a model function Θ^0 for the spins to compute $\mathbf{U}(P)$, since the exact function Θ is not known. This introduces an additional approximation to the wave function Ψ defined by (2.3). In (2.3), we have specified the product Λ of atomic orbitals, but Θ can, in principle, be obtained by solving the spin Hamiltonian of the type (1.2).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

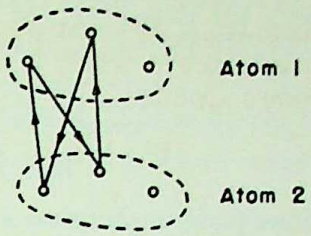


FIG. 12. An example of diagrams for arrays of many-electron atoms.

to group electrons belonging to a single atom and draw lines only between groups as is shown in Fig. 12. Of course *heavy* lines introduced in Figs. 10 and 11 representing interactions can connect electrons belonging to one atom.

The second remark concerns the case where two electrons occupy the same orbital with up and down spins. Here the overlap matrix $\mathbf{D}$ no longer has an inverse when the entire space G is considered. As we have discussed in Appendix B of paper I, however, there is a smaller space in which $\mathbf{D}^{-1}$ exists and our method can be used without any further difficulty. The smaller space is in fact a complete space of spin wave spectra for this case. We also do not need to consider $N!$ permutations but take one out of those generated from each other by simply permuting electrons in the same orbitals. For example, we need to consider only one of the diagrams in Fig. 13 when h_1 and h_2 are the same atomic orbital. Detailed discussion of such techniques will be found elsewhere[24] and therefore we shall not discuss it any further.

ACKNOWLEDGMENTS

In the course of developing the present work, the author has received great benefits from numerous communications with Dr. C. Herring. In fact, Dr. C. Herring has originally suggested the theorem described in the Appendix. Thanks are due to Dr. T. L. Gilbert for many discussions.

APPENDIX: EXPANSION OF $(1+\mathbf{X}[h(n)])^{-1}$

In this Appendix, we shall discuss the convergence of the expansion (3.14) or (3.10). First, it will be shown that, if the inequality

$$\epsilon=\sum_P |x_P[h(n)]|<1 \tag{A1}$$

is valid, the series (3.14) converges.

Both sides of (3.14) can be expanded in terms of $\tilde{\mathbf{U}}(P)$ and it will be found that

$$f_P[h(n)]=\delta_{EP}-(x[h(n)])_P+(x[h(n)]^2)_P-\cdots, \tag{A2}$$

where

$$(x[h(n)]^m)_P = \sum_{Q_1}\cdots\sum_{Q_{m-1}} x_{Q_1}[h]x_{Q_2Q_1^{-1}}[h]\cdots x_{PQ_{m-1}^{-1}}[h]. \tag{A3}$$

Taking the sum of absolute values of $(x[h(n)]^m)_P$, we find that

$$\sum_P |(x[h(n)]^m)_P| \leqslant [\sum_Q |x_Q[h(n)]|]^m=\epsilon^m. \tag{A4}$$

Inserting (A4) into (A2), we obtain

$$f_P[h(n)]\leqslant \delta_{EP}+\epsilon+\epsilon^2+\cdots=\delta_{EP}+\epsilon(1-\epsilon)^{-1}. \tag{A5}$$

and also

$$\sum_P |f_P[h(n)]| \leqslant (1-\epsilon)^{-1}. \tag{A6}$$

These results prove the original statement. If we truncate the series after the mth term of (3.14), errors in $f_P[h(n)]$ and $\mathbf{F}[h(n)]$ will both be smaller than $\epsilon^m(1-\epsilon)^{-1}$.

In the forthcoming paper, we shall introduce the theorem that the sum L_n of absolute values of the expansion coefficients $f_P[h(n)]$ is bounded as

$$[1+O(\Delta\Delta_0)]^{-n}\leq L_n\leq[1-O(\Delta\Delta_0)]^{-n}, \tag{A7}$$

where

$$L_n=\text{max on } h\text{'s of } \sum_P |f_P[h_1\cdots h_n|h_1\cdots h_n]|, \tag{A8}$$

$$\Delta=\text{max on } h \text{ of } \sum_{j\neq h} |\langle j|h\rangle|, \tag{A9}$$

$$\Delta_0=\text{max of } |\langle j|h\rangle|, \tag{A10}$$

and

$$O(\Delta\Delta_0)=\frac{\Delta\Delta_0}{1-\Delta}\left\{1+\frac{1}{4}\frac{4\Delta\Delta_0}{(1-\Delta)^2}+\frac{1.3}{4.6}\left[\frac{4\Delta\Delta_0}{(1-\Delta)^2}\right]^2+\cdots\right\}. \tag{A11}$$

Here we shall show briefly how this theorem can be proved. The expansion formula (2.23) of matrices S leads to

$$\mathbf{F}[\cdots h_n|\cdots h_n] = \mathbf{F}[\cdots h_nh_{n+1}|\cdots h_nh_{n+1}] + \sum_h{}^{(n+1)}\langle h|h_{n+1}\rangle \mathbf{F}[\cdots h_nh|\cdots h_nh_{n+1}], \tag{A12}$$

and

$$\mathbf{F}[\cdots h_{n+1}|\cdots h_n] = \langle h_n|h_{n+1}\rangle \mathbf{F}[\cdots h_{n+1}h_n|\cdots h_nh_{n+1}] + \sum_h{}^{(n+1)}\langle h|h_{n+1}\rangle \mathbf{F}[\cdots h_{n+1}h|\cdots h_nh_{n+1}], \tag{A13}$$

where $\cdots$ in the square brackets indicates $h_1h_2\cdots h_{n-1}$. When the matrices $\mathbf{F}$ in (A12) are expanded in terms of f_P and $\tilde{\mathbf{U}}(P)$, the equation is converted into a set of $N!$ equations for f_P as follows:

$$f_P[\cdots h_n|\cdots h_n] = f_P[\cdots h_nh_{n+1}|\cdots h_nh_{n+1}] + \sum_h{}^{(n+1)}\langle h|h_{n+1}\rangle f_P[\cdots h_nh|\cdots h_nh_{n+1}], \tag{A14}$$

[24] M. Kotani *et al.*, Ref. 18, also see T. Arai, J. Phys. Soc. Japan 18, 718 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

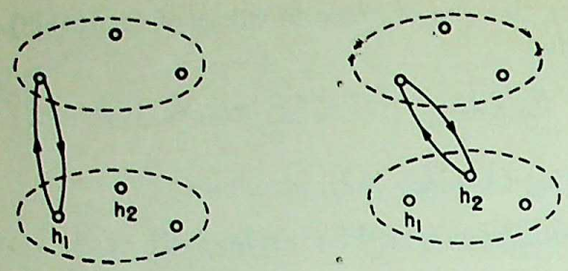


FIG. 13. An example of a case of two equivalent diagrams. Here electrons h_1 and h_2 occupy the same atomic orbital with up and down spins. We need to consider only one of them.

for all P. By taking absolute values of each term in (A14) and summing over $N!$ permutations P, we obtain that

$$\sum_P |f_P[\cdots h_n|\cdots h_n]|$$

$$\leqslant \sum_P |f_P[\cdots h_n h_{n+1}|\cdots h_n h_{n+1}]| + \sum_h{}^{(n+1)} |\langle h|h_{n+1}\rangle| \times \sum_P |f_P[\cdots h_n h|\cdots h_n h_{n+1}]| . \quad \text{(A15)}$$

Since the above inequality is valid for any h's, we find that

$$L_n \leq L_{n+1} + \Delta l_{n+1}{}^1, \quad \text{(A16)}$$

where

$$l_n{}^1 = \max \text{ on } h\text{'s of } \sum_P |f_P[\cdots h_n h|\cdots h_n h_{n+1}]| , \quad \text{(A17)}$$

and h is not involved among $h_1 \cdots h_n h_{n+1}$.

Similarly transposition of (A12) leads to

$$L_{n+1} \leq L_n + \Delta l_{n+1}{}^1, \quad \text{(A18)}$$

and, from (A13), we also find that

$$l_n{}^1 \leq \Delta_0 L_{n+1} + \Delta l_{n+1}{}^1. \quad \text{(A19)}$$

If the relation

$$l_{n+1}{}^1 \leq [O(\Delta\Delta_0)/\Delta] L_{n+1} \quad \text{(A20)}$$

is valid for any n, the present theorem will be proved. In fact, use of (A20) in (A16) and (A18) leads to

$$[1+O(\Delta\Delta_0)]^{-1} L_n \leq L_{n+1} \leq [1-O(\Delta\Delta_0)]^{-1} L_n \quad \text{(A21)}$$

for any n. If n becomes zero, the matrix $\mathbf{F}$ becomes $\mathbf{D}\mathbf{D}^{-1}=1$ and $L_0=1$. Inserting this result in (A21) when $n=0$, we find the desired bound for L_1. Repeated use of (A21), by increasing n one by one, will lead to the expression (A7).

The relation (A20) will be proved by an inductive unraveling process as follows. When $n=N-1$, Eq. (A19) becomes

$$l_{N-1}{}^1 \leq \Delta_0 L_N, \quad \text{(A22)}$$

since the second term on the right of (A13) vanishes. From (2.19), it is clear that $\mathbf{F}[\cdots h_{N-1}|\cdots h_{N-1}] = \mathbf{F}[\cdots h_{N-1}h_N|\cdots h_{N-1}h_N]$, and hence $L_{N-1}=L_N$. This implies that (A22) is written as

$$l_{N-1}{}^1 \leq \Delta_0 L_{N-1}. \quad \text{(A23)}$$

Since Δ_0 is smaller than $O(\Delta\Delta_0)/\Delta$, the above relation proves the inequality (A20) when $n=N-1$.

Let us assume that (A20) is valid for $n=n_0$. Use of (A20) in (A18) and (A19) yields

$$[1-O(\Delta\Delta_0)] L_{n_0+1} \leq L_{n_0} \quad \text{(A24)}$$

and

$$l_{n_0}{}^1 \leq [\Delta_0 + O(\Delta\Delta_0)] L_{n_0+1}. \quad \text{(A25)}$$

From (A24) and (A25), we obtain

$$l_{n_0}{}^1 \leq [\Delta_0 + O(\Delta\Delta_0)][1-O(\Delta\Delta_0)]^{-1} L_{n_0}. \quad \text{(A26)}$$

Inserting the expression (A11) for $O(\Delta\Delta_0)$ into (A26), we will find that the coefficient

$$[\Delta_0 + O(\Delta\Delta_0)][1-O(\Delta\Delta_0)]^{-1}$$

on the right of (A26) is equivalent to $O(\Delta\Delta_0)/\Delta$. Hence (A20) is also valid for $n=n_0-1$. This shows that (A20) is valid for any n. Thus the present theorem is proved.

Use of the inequality (A7) in (3.9) leads to

$$\epsilon = \sum_P |x_P[h]| < \Delta^2(1-O(\Delta\Delta_0))^{-1} + \Delta^3(1-O(\Delta\Delta_0))^{-2} + \Delta^4(1-O(\Delta\Delta_0))^{-3} + \cdots. \quad \text{(A27)}$$

Therefore, the inequality (A1) is satisfied when $n=1$ and the expansion (3.10) is valid as long as Δ is small as compared with one.

Use of the inequality (A7) in (3.13) will lead in the case $n=2$ to the same conclusion and it will be found that the series (3.12) converges. If, on the other hand, n increases, the upper bound $(1-O(\Delta\Delta_0))^{-n}$ and ϵ increase indefinitely and therefore the expansion (3.14) is no longer valid. However, it will be found that, in calculating an expectation value $\langle \mathbf{f} \rangle$ or $\langle \mathbf{g} \rangle$, the matrix $\mathbf{F}[h(n)]$ appears in the form $S_{R_n}\mathbf{F}[h(n)]$, which vanishes when n increases indefinitely since

$$S_{R_n}\mathbf{F}[h(n)] < \Delta^n(1-O(\Delta\Delta_0))^{-n} \to 0 \quad \text{(A28)}$$

as long as Δ is small as compared with unity. Here R_n is a cyclic permutation with which electrons $h_1, h_2, \cdots, h_{n-1}, h_n$ are shifted to orbitals $h_2, h_3, \cdots, h_n, h_1$ and S_{R_n} is an overlap of the type $\int R_n \Lambda^* \Lambda dv$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Thomas-Fermi Theory of the Atom as a Solution of the Density-Matrix Hierarchy*†

NANDOR L. BALAZS
State University of New York at Stony Brook, Stony Brook, New York

(Received 15 July 1963; revised manuscript received 16 December 1963)

In this paper we show that the Thomas-Fermi (TF) theory for a neutral atom is the zeroth-order solution of the full N-body problem if (a) the following assignments of smallness are made: $N^{-1/3}=\epsilon$, $e=e_0\epsilon^{3/2}$, $m=m_0/\epsilon^2$, $\epsilon\to 0$; (b) the singlet density matrix in the x representation, $\rho_1(x',x)$, is a fast oscillating function of the off-diagonal elements, of the type $\sin[(x'-x)/\epsilon]$; (c) the higher-order density matrices are determinants of the singlet density matrix as $\epsilon\to 0$. The higher-order approximations, however, cannot be obtained by a simple power-series expansion in ϵ, since the solutions contain ϵ in a nonanalytical fashion. Taking into account the exclusion principle to zeroth order in ϵ, and solving the equations of motion for the singlet density matrix to the next-higher-(second-) order approximation, we obtain the equations found by Kompaneets and Pavlovskii, and Baraff and Borowitz. These contain the Dirac and Weizsäcker corrections. Finally, we offer some conjectures about possible improvements of the approximation scheme.

1. INTRODUCTION

THE Thomas-Fermi (TF) model of the atom is the most successful attempt to account in simple terms for the characteristic features of complex atoms.

Many papers deal with the application of this theory to different systems, and with its heuristic improvements.[1] However, relatively little attention has been paid to the *systematic construction* of a statistical theory of complex atoms. The original derivation of the TF model was heuristic, and it is hard to tell what are the approximations involved, and how one can improve upon them. Such a more systematic treatment can be obtained in two different ways: expansion in terms of a small parameter, or a variational treatment.

There were three major efforts utilizing an expansion technique, all starting essentially from the Hartree-Fock (HF) scheme, expanding (1) in Planck's constant, $\hbar$[2]; (2) in inverse powers of the number of electrons (N) in the atom[3]; (3) in the off-diagonal elements of the density matrix.[4]

There are, however, two major difficulties associated with any of these attempts. First, it is not clear in what sense one should take h, $1/N$, or the magnitude of the off-diagonal elements of the density matrix as small. It cannot mean that in the limit when the chosen parameter tends to zero we obtain a rigorous solution, since if h goes to zero the exclusion principle must become inoperative, and if N goes to infinity the atom acquires an infinite radius. Clearly, a coupled limiting process is involved and one has to determine which parameters tend to zero and at what rate. Second, starting with the HF scheme the following objections are left unanswered. (A) One can show at most in this way that in some limit the TF model is as good as the HF scheme; actually it might turn out to be better. (B) We may indeed find that in some limit we obtain from the HF scheme the TF solutions; it also may happen that in this limit the HF scheme itself is a very poor approximation to the actual N-body problem. (C) If we intend to improve systematically the TF model starting with the HF scheme we have no assurance that the correction terms found are not of the same order of magnitude as the error committed by starting with the HF scheme in the first place.

Another very popular approach makes use of the fact that the HF scheme can be obtained as the Euler equations of a variational principle.[5] Write out the energy of the atom as a functional of the one-particle density matrix. Stipulate as auxiliary conditions that the one-particle density matrix or singlet density matrix normalized to N has a fixed trace and is idempotent. The Euler equations of this variational problem give the HF scheme. One is then tempted to go into this variational problem with a suitable trial function and determine its best form in the usual manner. However, a very serious trouble occurs (apart from the fact that the accuracy of the approximation cannot be assessed). The one-particle density matrix and hence the trial function has to obey an additional condition which cannot be given a simple mathematical form without solving the whole N-body problem: The singlet density matrix is to be of the form which can be obtained by integrating $N-1$ times the N-particle density matrix. While it is obvious that for the trial function this implies "smoothness" in some sense, it is by no means clear how to formalize this condition. Without satisfying it, however, the variational approximations obtained can be void of any physical meaning.

* Work supported by the National Science Foundation.

† Invited paper at the Washington meeting of the American Physical Society, April 1963.

[1] P. Gombas, *Die Statistische Theorie des Atoms und ihre Anwendungen* (Springer-Verlag, Vienna, 1949); *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. 36, p. 109.

[2] P. A. M. Dirac, Proc. Cambridge Phil. Soc. 26, 376 (1930); G. A. Baraff and S. Borowitz, Phys. Rev. 121, 1704 (1961).

[3] A. S. Kompaneets and E. S. Pavlovskii, Zh. Eksperim. i Teor. Fiz. 31, 427 (1956) [English transl.: Soviet Phys.—JETP 4, 328 (1957)].

[4] M. A. Naqui, Nucl. Phys. 10, 256 (1959).

[5] J. E. Mayer, J. Chem. Phys. 100, 1579 (1955); H. Koppe, Z. Physik 148, 135 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

For this reason we decided to investigate the possibility of an expansion in terms of a small parameter. Of course, it is possible that the wisest approach will eventually combine both methods. Throughout this paper we confine our interest to neutral atoms.

2. OUTLINE

Our plan will be as follows: First, we find the small parameter involved; then we show heuristically that in the limit when this parameter goes to zero the TF model is to be expected to give a rigorous solution of the N-body problem in question, and we assess the orders of magnitude of the different correction terms usually appended to the original TF theory (see Ref. 1). Then we start from first principles, set up the appropriate equations and symmetry conditions and show that indeed to zeroth order in this parameter we *rigorously* obtain the conventional TF theory. Next we shall show that the actual solutions are nonanalytical in the small parameter, hence the next approximation cannot be obtained by a simple power-series expansion. If, notwithstanding we attempt to do this in an inconsistent fashion, by improving the solutions of the differential equations to the next order, while satisfying the exclusion principle only to zeroth order, we obtain the usual exchange (Dirac) and inhomogeneity (Weizsäcker) correction. This will demonstrate, then, that these correction terms are probably not quite correct, and that for a consistent improvement over the original TF results it is necessary to devise an approximation method which can be used if the solution is nonanalytical in the small parameter.

3. THE SMALL PARAMETER

Characteristic Quantities

To construct the small parameter we shall first need the characteristic quantities involved in the TF theory. A short, dimensional analysis will accomplish this. Assume that the electron cloud around the nucleus is confined to a volume L^3; let this be accomplished entirely by an electrostatic potential U in cooperation with the exclusion principle. (This, of course, may not be possible, but if it is, the argument runs as follows.) Let there be N electrons confined; denote the maximum momentum at a given point inside L^3 by p_0, the mass and charge of an electron by m and e. Then, from Poisson's equation

$$U/L^2 \approx e^2 N/L^3, \tag{1}$$

since the density is $\approx N/L^3$. The condition that p_0 is the maximum momentum reads

$$p_0^2/2m \approx U. \tag{2}$$

The exclusion principle stipulates that each phase cell should contain at most one electron with a given spin; the condition that the atom be in its lowest energy state requires that no phase cell having momentum less than p_0 be empty. The two conditions together give

$$p_0^3/h^3 \approx N/L^3. \tag{3}$$

From (1) and (2) it follows that

$$U \approx Ne^2/L \approx p_0^2/2m, \tag{4}$$

while (2) and (3) give

$$p_0 L \approx hN^{1/3}, \quad L \approx a/N^{1/3}; \quad (a = \text{Bohr radius}).$$

Thus, we obtained the three characteristic quantities p_0, U, and L as functions of N. (Observe that the $1/N^{1/3}$ dependence is connected with the three-dimensional nature of the problem. One- and two-dimensional atoms, composed of charged sheets, or charged filaments, respectively, would have characteristic quantities with a different N dependence.)

Assessment of the TF Approximation

Suppose you have solved the TF differential equations. How can you tell whether the answer is, or is not, a correct solution of the full N-body problem? The problem is not trivial since the TF solutions do not furnish immediately the wave function or the density matrix of the N-body problem. There are two avenues open. First, the TF solution gives the number of electrons with given positions and momenta; from this we may construct in the well-known manner the singlet density matrix. In turn, from this we may construct the density matrix of the whole system, provided further assumptions are made about the relation of the singlet density matrix to the complete density matrix. Once this is obtained, we may verify how far this matrix obeys the equations of motion and the exclusion principle. This method has the advantage that both the satisfaction of the equations of motion and the exclusion principle can be tested. However, from the TF theory it is not immediately clear what are those further assumptions necessary to construct the density matrix from the singlet density matrix. For this reason we choose the second approach. This will test to a certain degree the satisfaction of the equations of motion, but not that of the exclusion principle. The TF theory is semiclassical theory, consequently it assumes that the local de Broglie wavelength varies slowly from point to point. Is this assumption satisfied everywhere by the solution? If we plot according to the TF theory the number of electrons contained in a spherical shell of radius r and unit-thickness around the nucleus, $n(r)4\pi r^2$, where $n(r)$ is the density at the point r, we find a function as shown in Fig. 1. For small values of r the function starts as $\sqrt{r}$, it has a single maximum at about $a/N^{1/3} = L$, where a is the Bohr radius and N is the number of electrons in the neutral atom; for large values of r the function goes to zero as $r^{-3/2}$. Associated with this density there is an electric potential energy $U(r)$ which goes as Ne^2/r if $r \ll L$, and as e^2a^3/r^4 for $r \gg L$. Finally we can define a local momentum

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

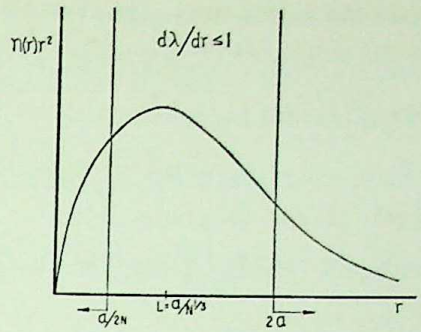


FIG. 1. Qualitative graph showing number of electrons on a shell of radius r as given by the Thomas-Fermi theory. Approximation good within region where the rate of change of the local de Broglie wavelength is small.

$p(r)=[2m(U-E)]^{1/2}$. The validity of the semiclassical approximation requires that $d\lambda\!\!\!^{-}/dr\ll 1$, with $\lambda\!\!\!^{-}=\hbar/p$. Write $d\lambda\!\!\!^{-}/dr=\lambda\!\!\!^{-}d\ln p/dr$; put approximately $\ln p\sim\frac{1}{2}\ln U$ +const with $U\sim r^{-n}$; $n=1$, $r\ll L$, $n=4$, $r\gg L$. This gives $\lambda\!\!\!^{-}n/2r\ll 1$ as our condition. The latter is certainly violated if $\frac{1}{2}n(\lambda\!\!\!^{-}/r)\equiv\frac{1}{2}n(\hbar/pr)$ is equal to 1. There are two roots to this equation: $a/2N$ and $2a$. Between then lies the maximum of the charge distribution (see Fig. 1). Hence we may say that for the majority of the electrons the TF solution is a good approximation to the Schrödinger equation.

Is the exclusion principle satisfied? Seemingly yes, since the solution was so constructed that one, or no, electrons of a given spin occupy each phase cell. However, this elementary form of the exclusion principle does not hold any longer if density gradients are present within the system and a more complicated condition has to be satisfied in its place. The derivation of this condition requires more preparation. In Sec. 7 we will show what this condition is, and how well the TF solutions satisfy it.

Construction of the Small Parameter

Now let us find the small parameter of the problem. Obviously we seek a limiting process such that in this limit the equations and the exclusion principle should be satisfied everywhere, while keeping the characteristic features of the solution unchanged. The usual solution is believed to be correct for $a/2N<r<2a$. The characteristic features of the atom represented by this solution are the radius of the atom L, and the energy per electron Ne^2/L. (According to the virial theorem for charged particles the kinetic and the potential energies, hence the total energy have the same order of magnitude.) Thus we seek a limiting process where a/N goes to zero, a goes to infinity, while L and Ne^2/L is being kept fixed. If we keep h fixed, this can only be accomplished by letting N go to infinity, e going to zero as $N^{-1/2}$, and m going to infinity as $N^{2/3}$, thus $N\to\infty$, $e=e_0N^{-1/2}$, $m=m^0N^{2/3}$, e_0, m_0 being the original values of e and m. In this limit then the equations are satisfied everywhere, and, as we shall show, so is the exclusion principle. Formally, we may consider h as variable, instead of m. Then the same result can be obtained by letting h go to zero as $N^{-1/3}$, i.e., $N\to\infty$, $e=e_0N^{-1/2}$, $\hbar=\hbar_0N^{-1/3}$. Since we are employing a limiting process in which N tends to infinity, the question arises as to the computation of extensive quantities. The rule is as follows. First, compute the value of the associated intensive quantity in the limit as N tends to infinity. (E.g., if you are interested in the total energy, compute first the energy per particle as N tends to infinity.) Next, multiply this quantity by the actual number of particles.

Finally, we would like to stress that this selection of a small parameter was specifically designed for the purposes of the TF problem in such a manner as to make the semiclassical approximation valid for the entire atom in the zeroth order. For this reason it has less generality than the usual method where one treats the difference between the actual and the self-consistent potential as small; on the other hand, because of the more specific nature of the choice it tells us more about the approximation.

Orders of Magnitude of the Correction Terms

The usual exchange correction to the energy per electron is given by $\approx e^2p_0/h$, i.e., the Coulomb interaction between two electrons at a de Broglie wavelength h/p_0 apart. The ratio of this quantity to p_0^2/m is of the order $1/N^{2/3}=\epsilon^2$. The inhomogeneity correction to the energy per electron is given by $\approx e^2a[\text{grad}\log n]^2$; its ratio to p_0^2/m is again ϵ^2. Thus, if we accept ϵ as a relevant small parameter, these terms are correction terms of second-order smallness. The Fermi-Amaldi correction, replacing N with $N-1$, gives obviously a correction of order $1/N=\epsilon^3$, a term of third-order smallness. Finally, the correlation correction, introduced by Gombás, $\sim e^2n^{1/3}/(an^{1/3}+10^{-1})$ is of the order $\epsilon^4p_0^2/m$.

4. THE BASIC EQUATIONS

In this section we derive the basic equations of the theory, the equations of motion for the reduced density matrices. We shall work with density matrices for the following reason. The basic aim is to set up expressions in which the coordinates of one or two electrons appear only, the coordinates of the other electrons being integrated out. However, the state of one electron interacting with the others cannot be described by a wave function (not being a pure state) while it can be described by a density matrix; hence the latter is the proper tool.

Envisage a neutral atom with N electrons, and a fixed nucleus with charge number N. The Hamiltonian of the system is given by

$$H=\sum_{r=1}^{N}(K_r+V_r)+\sum_{1\leqslant i<j\leqslant N}\Phi(x_i-x_j),$$

$$K_r=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_r\partial x_r};\quad \Phi(x)=e^2/|x|;\qquad (5)$$

$$V_r=-N\Phi(x_r).$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The state of the system is described by the antisymmetrical wave function $\Psi(x_1,\cdots,x_N,t)$. (In principle the spin coordinates should also be included; however, the absence of spin variables in the Hamiltonian has the effect that the existence of the spin coordinates changes only certain factors in the equations. These can be put in any time, practically by inspection. Hence, for simplicity, we omit them throughout the paper, and shall point out when necessary which terms acquire different factors.) The density matrix for the whole atom in a pure state is defined as

$$\rho_N(x_1',\cdots,x_N';\,x_1,\cdots,x_N,t) = \Psi^*(x_1',\cdots,x_N',t)\Psi(x_1,\cdots,x_N,t). \quad (6)$$

Differentiate (6) with respect to t and express the time derivative of Ψ through the Schrödinger equation. We obtain

$$i\hbar\partial\rho_N/\partial t=(H-H')\rho_N, \quad (7)$$

where H' acts only on the x' variables. (This equation holds true even if ρ_N does *not* represent a pure state.)

Define now the reduced density matrices ρ_s by

$$\rho_s(x_1',\cdots,x_s',x_1,\cdots,x_s,t)=\int d^3x_{s+1}\cdots \times\int d^3x_N\rho_N(x_1',\cdots,x_s',x_{s+1},\cdots,x_N;\,x_1,\cdots,x_s,\cdots,x_{s+1},\cdots,x_N,t). \quad (8)$$

Contrary to the usual definition we do not include any combinatorial factors in the definition of the ρ_s. Consequently, all the density matrices $\rho_s(s=1, 2, \cdots, N)$ are normalized to unity. The reason for this departure is due to the fact that an N-dependent normalization would make a limiting process $N\to\infty$ more unpleasant. In Eq. (7) put $x_{s+1}'=x_{s+1}$, $x_{s+2}'=x_{s+2}$, $\cdots$, $x_N'=x_N$, and integrate over these variables. We get

$$i\hbar\frac{\partial\rho_s}{\partial t}=\sum_{r=1}^{s}[(K_r-K_r')+(V_r-V_r')]\rho_s + \sum_{1\leqslant i<j\leqslant s}[\Phi(x_i-x_j)-\Phi(x_i'-x_j')]\rho_s+(N-s) \times\sum_{1\leqslant i\leqslant s}\int d^3x_{s+1}[\Phi(x_i-x_{s+1})-\Phi(x_i'-x_{s+1})] \times\rho_{s+1}(x_1',\cdots,x_s',x_{s+1};\,x_1,\cdots,x_s,x_{s+1}). \quad (9)$$

If the atom is in a stationary state, the time derivative is equal to zero. The exclusion principle is incorporated in the symmetry conditions on the ρ_s; an interchange of any two primed coordinates, reproduces the function with the opposite sign. (Since we shall discuss stationary states we may choose, if we so desire, all density matrices real.)

Define now the new dimensionless variables

$$y_k=(x_k'+x_k)/2L,\ z_k=(x_k'-x_k)/L,\ (k=1,\ 2,\ \cdots,\ N),$$

and the new dimensionless functions of these variables

$$\rho_s(x_1',\cdots,x_s';\,x_1,\cdots,x_s)=R_s(y_1,\cdots,y_s;\,z_1,\cdots,z_s)/L^{3s}.$$

Equation (8) becomes

$$R_s(y_1,\cdots,y_s;\,z_1,\cdots,z_s) =\int d^3y_{s+1}R_{s+1}(y_1,\cdots,y_s,y_{s+1};\,z_1,\cdots,z_s,0).$$

Write $\Phi(x)=e^2\phi/L$ and observe that $K_r(x)-K_r(x')=h^2/mL^2(\partial^2/\partial y_r\partial z_r)$. Substitute all these back in Eq. (9), and divide by Ne^2/L; we obtain (putting $\partial\rho_s/\partial t=0$)

$$\sum_{1\leqslant i\leqslant s}\{(a/LN)(\partial^2/\partial y_i\partial z_i) +[\phi(y_i+z_i/2)-\phi(y_i-z_i/2)]\}R_s -1/N\sum_{1\leqslant i<j\leqslant s}\{\phi[y_i-y_j+(z_i-z_j)/2] -\phi[y_i-y_j-(z_i-z_j)/2]\}R_s-(1-s/N) \times\sum_{1\leqslant i\leqslant s}\int d^3y_{s+1}[\phi(y_i-y_{s+1}+z_i/2) -\phi(y_i-y_{s+1}-z_i/2)]R_{s+1}(y_1,\cdots,y_s,y_{s+1};\,z_1,\cdots,z_s,0)=0.$$

Introduce $N=\epsilon^{-3}$ and $L/a=\epsilon$ according to the previous section. The first term acquires the factor ϵ^2. Absorb one ϵ in the new variable $\zeta=z/\epsilon$, and divide with the other ϵ. This way finally we obtain the set of equations, ϵ^{-3} in number, we shall work with:

$$\sum_{1\leqslant i\leqslant s}\left\{\frac{\partial^2}{\partial y_i\partial\zeta_i}+(1/\epsilon)[\phi(y_i+\epsilon\zeta_i/2)-\phi(y_i-\epsilon\zeta_i/2)]\right\} \times R_s(y_1,\cdots,y_s;\,\epsilon\zeta_1,\cdots,\epsilon\zeta_s|\,\epsilon) -\epsilon^3\sum_{1\leqslant i<j\leqslant s}(1/\epsilon)\{\phi[y_i-y_j+\epsilon(\zeta_i-\zeta_j)/2] -\phi[y_i-y_j-\epsilon(\zeta_i-\zeta_j)/2]\}R_s(y_1,\cdots,y_s;\, \epsilon\zeta_1,\cdots,\epsilon\zeta_s|\,\epsilon)-(1-\epsilon^3s)\sum_{1\leqslant i\leqslant s}\int d^3y_{s+1}(1/\epsilon) \times[\phi(y_i-y_{s+1}+\epsilon\zeta_i/2)-\phi(y_i-y_{s+1}-\epsilon\zeta_i/2)] \times R_{s+1}(y_1,\cdots,y_s,y_{s+1};\,\epsilon\zeta_1,\cdots,\epsilon\zeta_s,0|\,\epsilon)=0, \quad (s=1,\ 2,\ \cdots,\ 1/\epsilon^3). \quad (10)$$

Here we have explicitly indicated the fact that the R functions now can depend on ϵ in two different ways; first, due to the change of variables from z to ζ, and second, on account of the fact that originally ρ_s, and hence R_s, involved $N(=1/\epsilon^3)$ parametrically (since they were obtained from ρ_N or R_N by integration). We

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

rewrite Eq. (10) in a more transparent form. Introduce the difference operator D_r; it acts on the argument y_r of a function F to which it is applied;

$$D_rF(y_1,y_2,\cdots,y_r,\cdots) \\ -(1/\epsilon)[F(y_1,y_2,\cdots,y_r+\epsilon\zeta_r/2,\cdots) \\ -F(y_1,y_2,\cdots,y_r-\epsilon\zeta_r/2,\cdots)].$$

Thus in the limit $\epsilon\to 0$ it gives $\zeta_r\cdot\partial F/\partial y$. Equation (10) reads now as

$$\sum_{1\leqslant i\leqslant s}\left[\frac{\partial^2}{\partial y_i\partial\zeta_i}+D_i\phi(y_i)\right]f_s \\ -\epsilon^3\sum_{1\leqslant i<j\leqslant s}[(D_i+D_j)\phi(y_i-y_j)]f_s \\ -(1-\epsilon^3 s)\sum_{1\leqslant i\leqslant s}\int d^3y_{s+1}\int d^3\zeta_{s+1}\delta(\zeta_{s+1}) \\ \times[(D_i+D_{s+1})\phi(y_i-y_{s+1})]f_{s+1}=0. \quad (11)$$

For sake of symmetry we introduced the delta function $\delta(\zeta_{s+1})$ and an additional integration over ζ_{s+1}. The functions f_s are defined as:

$$R_s(y_1,\cdots,y_s;\,\epsilon\zeta_1,\cdots,\epsilon\zeta_s|\,\epsilon)=f_s(y_1,\cdots,y_s;\,\zeta_1,\cdots,\zeta_s|\,\epsilon).$$

These equations must be supplemented by conditions on the f_s functions. These are the recursion relations:

$$f_s(y_1,\cdots,y_s;\,\zeta_1,\cdots,\zeta_s|\,\epsilon)=\int d^3y_{s+1}f_{s+1} \\ \times(y_1,\cdots,y_s,y_{s+1};\,\zeta_1,\cdots,\zeta_s,0|\,\epsilon); \quad (12)$$

the normalization conditions:

$$\int d^3y_1\cdots\int d^3y_s f_s(y_1,\cdots,y_s;\,0,\cdots,0|\,\epsilon)=1; \quad (13)$$

the symmetry conditions:

$$f_s\left(\frac{x_1'+x_1}{2L},\cdots,\frac{x_s'+x_s}{2L};\frac{x_1'-x_1}{\epsilon L},\cdots,\frac{x_s'-x_s}{\epsilon L}\middle|\,\epsilon\right) \quad (14)$$

must change its sign upon interchange of x_k and x_r, or x_k' and x_r', $k\neq r$.

We wrote out the latter in terms of the variables x, x', to simplify the conditions. Equations (12) and (13) are not independent. If we define $f_0=1$, the normalization conditions are the consequence of (12) for $s=0$. Hence, in the future we need not mention (13) explicitly.

Then our aim is to solve the equation-system (11) for the set of f_s functions satisfying conditions (12), (13), and (14) in the limit $\epsilon\to 0$; moreover, if possible, to devise a scheme of successive approximations for small, nonzero ϵ. In addition we may require that the solution should give the lowest possible value for the mean energy per electron. In that case we seek the solution for the lowest energy state of the neutral atom. If the atom is at temperature $T=0$, we must require that ρ_N, or the associated f_N should represent a pure state as indicated by (6). [For nonzero temperatures all equations remain unchanged except for (6).] The purpose of the new variables y, z, and ζ is the following: Envisage $\rho_1(x',x)$ defined at the lattice points in the (x',x) plane, then the y, z labeling of these points corresponds to a 45° rotation, and a change in scales (see Fig. 2); y labels a distance along lines parallel to the diagonal, z the perpendicular distance from the diagonal. Since our limiting process shifts the problem in the semiclassical region, *the off-diagonal elements must decrease in size and oscillate fast*; thus only small values of z become of importance. For this reason we introduce the new variable $\zeta=z/\epsilon$, which amounts to introducing a new unit of length along the z axis, which is ϵ times the previous one, using thereby a yardstick more adapted to the problem. This is also evident from the fact that in terms of these variables the coefficients in the equation-system (11) *contain positive powers of* ϵ, *and as* $\epsilon\to 0$ *the coefficients become independent of* ϵ. (At this stage one could take the Fourier transform with respect to z or ζ, a natural step to do for oscillatory functions; this introduces the so-called Wigner functions. However, the general treatment seems to be simpler without this step, and consequently we shall formulate the theory in terms of the f functions, and will use Fourier transforms only as a tool to solve the resulting equations.)

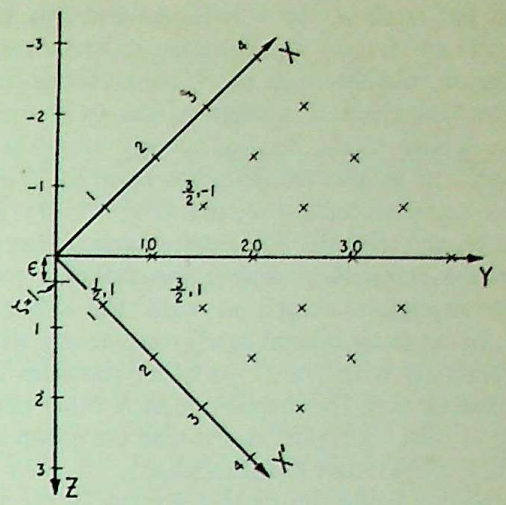


FIG. 2. Relation between old (x',x) and new (y,z) labeling of elements of singlet density matrix $\rho_1(x',x)$.

5. CHARACTERISTIC FEATURES OF THE EQUATION SYSTEM AND THE SUPPLEMENTARY CONDITIONS; NONANALYTICITY DIFFICULTIES

The solution of (11) subject to conditions (12), (13), and (14) is equivalent to the full, correct solution of the problem; in order to obtain an approximate solution we must study the characteristic features of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

problem for small ϵ. As ϵ tends to zero, in (11) all coefficients of f_s and f_{s+1} become independent of ϵ; moreover, we observe that there are no terms linear in ϵ, and two of the terms proportional to ϵ^3 disappear entirely. These terms have a simply physical interpretation. The set of s particles described by f_s interact with the external field [second term of (11)], with each other, and with the $N-s$ other particles; as ϵ tends to zero the interaction *within* the set of s particles becomes negligible compared with the other interactions, being proportional to ϵ^3; consistently with this the interaction with the $N-s$ other particles can be approximated with the interaction of N other particles. Because of this, the s particles within the group behave as if they would move independently of each other, their behavior being determined by the interaction of each with the particles outside the group. Thus we expect that up to and including order ϵ^2 the f_s functions should be expressible in terms of f_1. Thus the structure of the Eqs. (11) suggests that (a) using the variables y and ζ, f_s will depend only on positive powers of ϵ (this, of course, was just reason to introduce these variables); (b) there shall be no terms linear in ϵ; (c) up to ϵ^2, f_s shall be expressible in terms of f_1. Introduce now functions $f_s^{(k)}(y_1,\cdots,y_s,\zeta_1,\cdots,\zeta_s)$ *independent* of ϵ, where k denotes the order of the approximation. Then our expectations can be formulated in the expression

$$f_s(y_1,\cdots,y_s;\zeta_1,\cdots,\zeta_s|\,\epsilon)=\sum_{k=0}\epsilon^k f_s^{(k)}, \qquad (15)$$

with $f_s^{(0)}=F[f_1^{(0)}]$, $f_s^{(1)}\equiv 0$, $f_s^{(2)}=G[f_1^{(2)}]$; F and G are unknown functionals, independent of ϵ.

The following questions are to be faced: (a) Does (11) have solutions of this form at all; (b) if so, are these solutions compatible with the supplementary conditions; (c) if so, how to determine F and G together with $f_1^{(0)}$, $f_1^{(2)}$ from the equations and supplementary conditions? The difficulty we encounter in answering these questions is due to the fact that the introduction of the small parameter gives no clue as to the construction of the functions F and G, while the existence of the solutions will depend on this construction. Omitting in (11) the term containing the double sum on i and j, which is proportional to ϵ^3, we see that the equation contains only a sum of operators each acting on one set of y, ζ variables with a given index. This, then suggests that the functionals should simply be the product of their argument functions each carrying variables with a given index. Indeed, it is easy to show that one can find such solutions to (11) [see next section, Eq. (27)], however, they violate the symmetry conditions (4) for $s>1$, since they correspond to solutions which obey Boltzmann statistics.[6] It is natural that the symmetry condition will cause trouble if we use the variables y, ζ. In term of these variables it is *impossible to express the exclusion principle in an ϵ-independent way*. Thus the variables most suited for the discussion of the equations of motion are not so useful when it comes to express the conditions implied by the exclusion principle. Indeed the command: exchange x_1' and x_2', keep x_1, x_2 fixed, amounts to replacing y_1 with $\frac{1}{2}(y_1+y_2)+\frac{1}{4}\epsilon(\zeta_1-\zeta_2)$; y_2 with $\frac{1}{2}(y_1+y_2)-\frac{1}{4}\epsilon(\zeta_1-\zeta_2)$; ζ_1 with $(y_2-y_1)/\epsilon+\frac{1}{2}(\zeta_1+\zeta_2)$; ζ_2 with $-(y_2-y_1)/\epsilon+\frac{1}{2}(\zeta_1+\zeta_2)$. This explicit appearance of ϵ casts doubt on the possibility of expressing $f_s(|\,\epsilon)$ as a power series in ϵ with only positive powers, if the symmetry condition is to be satisfied. Indeed, for $s=2$ the symmetry condition reads

$$\begin{aligned}f_2(y_1,y_2;\zeta_1,\zeta_2|\,\epsilon)&\\ =-f_2[&\tfrac{1}{2}(y_1+y_2)+(\epsilon/4)(\zeta_1-\zeta_2),\ \tfrac{1}{2}(y_1+y_2)\\ &-(\epsilon/4)(\zeta_1-\zeta_2);\ (1/\epsilon)(y_2-y_1)+\tfrac{1}{2}(\zeta_1+\zeta_2),\\ &-(1/\epsilon)(y_2-y_1)+\tfrac{1}{2}(\zeta_1+\zeta_2)]. \qquad (16)\end{aligned}$$

Now we see why f_2 cannot be a power series involving only positive powers of ϵ, with coefficients which are functions of y and ζ only. f_2 will be oscillatory in ζ_1, and ζ_2, i.e., in the third and fourth variables. On account of (16), this means that it must also be oscillatory in $(y_1-y_2)/\epsilon$ thus it will have to contain, in general, arbitrarily high negative powers of ϵ. This does not preclude the possibility that there should be a solution of $\epsilon\to 0$; it only asserts that the ϵ dependence of the solution cannot be analytic around $\epsilon=0$ if it also satisfies the exclusion principle. This squashes our hopes of finding a solution of the form (15), and we must go about the solution differently. First, let us see whether there *is* a solution for $\epsilon=0$; next, investigate the possibility of an approximation scheme which is not a simple power-series expansion in ϵ. In the remainder of the paper we will discuss the $\epsilon=0$ approximation and only touch upon the more difficult problem of the approximation scheme.

6. THE ZEROTH-ORDER APPROXIMATION

The Reduction of the Hierarchy

For $\epsilon\to 0$, $D_r\to\zeta_r(\partial/\partial y_r)$, and the equations read

$$\left\{\frac{\partial^2}{\partial y_1\partial\zeta_1}+\zeta_1\frac{\partial\phi(y_1)}{\partial y_1}\right\}f_1(y_1,\zeta_1)$$
$$-\int d^3y\left[\zeta_1\frac{\partial}{\partial y_1}\phi(y_1-y_2)\right]f_2(y_1,y_2;\zeta_1,0)=0, \qquad (17a)$$

$$\sum_{1\leqslant i\leqslant s}\left\{\frac{\partial^2}{\partial y_i\partial\zeta_i}+\zeta_i\frac{\partial\phi(y_i)}{\partial y_i}\right\}f_s(y_1,\cdots,y_s;\zeta_1;\cdots,\zeta_s)$$
$$\times\sum_{1\leqslant i\leqslant s}\int d^3y_{s+1}\left[\zeta_i\frac{\partial\phi(y_i-y_{s+1})}{\partial y_i}\right]$$
$$\times f_{s+1}(y_1,\cdots,y_s,y_{s+1};\zeta_1,\cdots,\zeta_s,0)=0,$$
$$(s=2,3,\cdots); \qquad (17b)$$

[6] It is rather interesting that such a solution should still exist, since in constructing the small parameter ϵ we made already explicit use of the exclusion principle.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

with

$$f_s(y_1,\cdots y_s;\zeta_1,\cdots,\zeta_s)=\int d^3y_{s+1}f_{s+1}$$

$$\times(y_1,\cdots,y_s,y_{s+1};\zeta_1,\cdots,\zeta_s,0); \quad (18a)$$

where

$$f_s\left(\frac{x_1'+x_1}{2L},\cdots,\frac{x_s'+x_s}{2L};\frac{x_1'-x_1}{\epsilon L},\cdots,\frac{x_s'-x_s}{\epsilon L}\right) \quad (18b)$$

must change its sign if we interchange any two primed variables, x_i', x_j', say $(i=1,\cdots,s;\ j=1,\cdots,s;\ i\neq j)$, or any two unprimed variables, x_i, x_j, say; all f_s must be normalized to unity, the f_s must describe a pure state, the energy must be a minimum.

Now we construct the functional F, guided by the Hartree-Fock scheme. To do this we temporarily return to the density matrices ρ_s. Let $f_1^{(0)}(y_1,\zeta_1)$ be independent of ϵ, *qua* function of y and ζ; a $\rho_1(x',x)$ is now defined by

$$L^3\rho_1(x_1',x_1)=f_1^{(0)}\left(\frac{x_1'+x_1}{2L},\frac{x_1'-x_1}{\epsilon L}\right).$$

ρ_1, of course, will depend on ϵ in a complicated way. We proceed to f_s and ρ_s, $(s>1)$ in a similar manner. With $f_s(y_1,\cdots,y_s;\zeta_1,\cdots,\zeta_s|\,\epsilon)$ a $\rho_s(x_1',\cdots,x_s';x_1,\cdots,x_s)$ is associated by

$$f_s(y_1,\cdots,y_s;\zeta_1,\cdots,\zeta_s|\,\epsilon)=L^{3s}\rho_s(x_1',\cdots,x_s';x_1,\cdots,x_s),$$

where

$$x_k'/L=y_k+\tfrac{1}{2}\epsilon\zeta_k,\quad x_k=y_k-\tfrac{1}{2}\epsilon\zeta_k.$$

The functional relation is now obtained as follows:

First, construct ρ_s as a determinant,

$$\rho_s(x_1',\cdots,x_s';x_1,\cdots,x_s)$$

$$=\frac{N^s(N-s)!}{N!}\begin{vmatrix}\rho_1(x_1',x_1)&\cdots&\rho_1(x_1',x_s)\\ \vdots&&\vdots\\ \rho_1(x_s',x_1)&\cdots&\rho_1(x_s',x_s)\end{vmatrix}. \quad (19)$$

Next, express all the x's through y and ζ, and put $N=\epsilon^{-3}$. Finally, let $\epsilon\to 0$ wherever it still appears, after everything is expressed in terms of y and ζ (not x!). The resulting $f_s^{(0)}$ will be our zeroth approximation. (The appearance of the N^s factor in addition to the more usual combinational term is due to the fact that here ρ_1 is normalized to unity and not to N. This way, as $N\to\infty$ the combinatorial factor goes to one.) These functions should satisfy the supplementary conditions and the equations of motion. Let us start with the supplementary conditions. The recursion relation between ρ_2 and ρ_1 reads as

$$\int\rho_2(x_1',x_2;x_1,x_2)d^3x_2=\rho_1(x_1',x_1).$$

According to (19), ρ_2 is given by

$$\rho_2(x_1',x_2';x_1,x_2)=[N/(N-1)][\rho_1(x_1',x_1)\rho_1(x_2',x_2)-\rho_1(x_1',x_2)\rho_1(x_2',x_1)].$$

Substituting this in the previous equation we obtain the integral condition

$$N\int d^3x''\rho_1(x',x'')\rho_1(x'',x)=\rho_1(x',x). \quad (20)$$

Now we show that if (20) is fulfilled, the other recursion relations for $s\geqslant 2$ are also satisfied. Take ρ_s with $x_s'=x_s$ and expand the determinant according to elements in the sth column; multiply the coefficients into the last row and integrate over x_s. The integration can be performed using (20) and we indeed obtain ρ_{s-1} as defined by (19). This shows that all recursion relations are fulfilled if (20) is satisfied. Let us demonstrate now that if (20) is satisfied the system is in a pure state, i.e., it is at the temperature zero. The system is in a pure state if ρ_N is idempotent, i.e., if

$$\int d^3x_1''\cdots d^3x_N''\rho_N(x_1',\cdots,x_N';x_1'',\cdots,x_N'')\times\rho_N(x_1'',\cdots,x_N'';x_1,\cdots,x_N)=\rho_N(x_1',\cdots,x_N';x_1,\cdots,x_N).$$

Again, we can perform the integrations using (20) and we find that ρ_N *is* idempotent. Thus, the only conditions not yet satisfied through (20) are the normalization condition on f_1, and the minimal condition on the energy. We now express (20) and the normalization condition in terms of $f_1^{(0)}$. Substitute in (20) $\rho_1=f_1^{(0)}(y,\zeta)/L^3$ with $x'/L=y+\frac{1}{2}\epsilon\zeta$, $x/L=y-\frac{1}{2}\epsilon\zeta$ and put $N=\epsilon^{-3}$. Change the integration variables from x'' to s by $(x''-x)/\epsilon L=s$; one gets

$$\int d^3s f_1^{(0)}(y+\tfrac{1}{2}\epsilon s,\zeta-s)f_1^{(0)}[y-\tfrac{1}{2}\epsilon(\zeta-s),s]=f_1^{(0)}(y,\zeta). \quad (21)$$

Letting $\epsilon\to 0$, we obtain

$$\int d^3s f_1^{(0)}(y,\zeta-s)f_1^{(0)}(y_1,s)=f_1^{(0)}(y,\zeta). \quad (22)$$

The normalization condition reads

$$\int f_1^{(0)}(y,0)d^3y=1. \quad (23)$$

Thus, our problem is reduced to show that the so constructed $f_s^{(0)}$ functions satisfy (17), (22), (23), and the minimal condition on the energy. We shall accomplish this in two steps. First, we show that the $f_s^{(0)}$ functions for $s>1$ satisfy (17b) if $f_1^{(0)}$ satisfies (17a); next, we show that the set of equations (17b), (22), and (23) for

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$f_1^{(0)}$, coupled with the minimum condition is nothing else but the TF theory. To accomplish the first step we need the $f_s^{(0)}$ functions in terms of $f_1^{(0)}$. Let us start with $f_2^{(0)}$. Before going to the limit $\epsilon=0$, we get from (19)

$$f_2(y_1,y_2;\zeta_1,\zeta_2|\epsilon)=\frac{1}{1-\epsilon^3}\Big\{f_1^{(0)}(y_1,\zeta_1)f_1^{(0)}(y_2,\zeta_2)$$
$$-f_1^{(0)}\Big(\frac{1}{2}(y_1+y_2)+\frac{\epsilon}{4}(\zeta_1-\zeta_2),\ \frac{y_1-y_2}{\epsilon}+\frac{1}{2}(\zeta_1+\zeta_2)\Big)$$
$$\times f_1^{(0)}\Big(\frac{y_1+y_2}{2}-\frac{\epsilon}{4}(\zeta_1-\zeta_2);$$
$$-\frac{(y_1-y_2)}{\epsilon}+\frac{1}{2}(\zeta_1+\zeta_2)\Big)\Big\}. \quad (24)$$

In the semiclassical region the off-diagonal elements of ρ_1 decrease in size and oscillate fast. In terms of f_1 this means that f_1 for a fixed value of y must tend to zero and oscillate fast as ζ tends to infinity. Thus we seek solutions for which $f_1(y,\infty)=0$. Now let ϵ tend to zero. The second term is then always zero, unless $y_1=y_2$ in which case it has the *finite* value

$$f_1^{(0)}[y_1,\tfrac{1}{2}(\zeta_1+\zeta_2)]f_1^{(0)}[y_2,\tfrac{1}{2}(\zeta_1+\zeta_2)].$$

Hence,

$$f_2^{(0)}(y_1,y_2;\zeta_1,\zeta_2)\underset{\epsilon=0}{=}f_1^{(0)}(y_1,\zeta_1)f_1^{(0)}(y_2,\zeta_2)$$
$$-\delta_{y_1-y_2,0}f_1^{(0)}[y_1,\tfrac{1}{2}(\zeta_1+\zeta_2)]f_1^{(0)}[y_2,\tfrac{1}{2}(\zeta_1+\zeta_2)], \quad (25)$$

where $\delta_{k,0}$ is a Kronecker δ and not a Dirac δ. [We must add the stipulation $\epsilon\to 0$ since the symmetry properties required by the exclusion principle are expressed through the variables x', x and their relation to y, ζ involves ϵ. This makes ϵ appear each time explicitly in the right-hand side of (25) whenever we interchange x' and x.] This result can be easily generalized for the higher $f_s^{(0)}$ functions. The rule is as follows. Think of the expanded determinant in (19); each term in the sum will contain $s\rho_1$ factors; of these, some ρ_1 factors will have the same index on both the primed and unprimed x variables, e.g., $\rho_1(x_k',x_k)$; the others will carry variables where the two indices differ, e.g., $\rho_1(x_k',x_l)$. To get the corresponding $f_s^{(0)}$, replace each ρ_1 where the indices are the same, say k, by $f_1^{(0)}(y_k,\zeta_k)$, and replace each $\rho_1(x_k',x_l)$, where the indices on the two variables are different (k and l, say) by $\delta_{y_k-y_l,0}f_1^{(0)}\times(y_k,\frac{1}{2}(\zeta_k+\zeta_l))$. Now we are prepared to solve (17) in the limit $\epsilon\to 0$. Write (17a) as

$$A_1f_1(1)+J_{1,2}[f_2(1,2)]=0 \quad (26)$$

where A_1 is the operator $\partial^2/\partial y_1\partial\zeta_1+\zeta_1(\partial\phi/\partial y_1)$, and $J_{1,2}$ is the operator $(-1)\int d^3y_2\int d^3\zeta_2\delta(\zeta_2)\zeta_1[\partial\phi(y_1-y_2)/\partial y_1]$. The arguments in the f functions are indicated by the index of the argument, e.g.,

$$f_2(1,2)=f_2(y_1,y_2;\zeta_1,\zeta_2).$$

The set (17b) reads then

$$\sum_{r=1}^{s}A_rf_s(1,\cdots,s)$$
$$+\sum_{r=1}^{s}J_{r,s+1}[f_{s+1}(1,\cdots,s+1)]=0. \quad (27)$$

We first show that (27) is satisfied if f_s is a product of $f_1^{(0)}$'s, provided $f_1^{(0)}$ satisfies (26). [By (25) f_s *is* this product if all y's are different.] Then $f_s(1,\cdots,s)=f_1^{(0)}(1)\cdots f_1^{(0)}(s)$, and $J_{r,s+1}[f_{s+1}(1,\cdots,s+1)]=f_1^{(0)}(1)\cdots f_1^{(0)}(s)J_{r,s+1}[f_1^{(0)}(s+1)]$. Equations (26) and (27) become

$$A_1f_1^{(0)}(1)+f_1^{(0)}(1)J_{1,2}[f_1^{(0)}(2)]=0; \quad (28)$$

$$1f_1^{(0)}(2)\cdots f_1^{(0)}(s)\{A_1f_1^{(0)}(1)$$
$$+f_1^{(0)}(1)J_{1,s+1}[f_1^{(0)}(s+1)]\}$$
$$+f_1^{(0)}(1)1f_1^{(0)}(3)\cdots f_1^{(0)}(s)\{A_2f_1^{(0)}(2)$$
$$+f_1^{(0)}(2)J_{2,s+1}[f_1^{(0)}(s+1)]\}$$
$$+\text{same for index } 3+\cdots$$
$$+\text{same for index } s=0. \quad (29)$$

Thus, if $f_1^{(0)}(1)$ satisfies (28), then (29) is automatically satisfied for any s. If the y's are not all different, we proceed as follows. First, we observe that in (26) the second term in f_2 [see (25)] does contribute nothing, since we integrate a function, which is zero everywhere except at a point, where it is finite.[7] Next, we observe that for the same reason the integral in (27) contributes nothing as long as in the integrand there is a Kronecker δ present which carries a y_{s+1}. All such terms must carry a factor $f_1^{(0)}(s+1)$ according to the construction; thus they must be $\rho_1(x_{s+1}',x_{s+1})$ times the minor of $\rho_1(x_{s+1}',x_{s+1})$ in the determinant. Consequently, *under the operator* $J_{r,s+1}$ *we can write* $f_{s+1}(1,\cdots,s+1)=f_s(1,\cdots,s)f_1^{(0)}(s+1)$, and (27) becomes

$$\sum_{r=1}^{s}A_rf_s(1,\cdots,s)+f_s(1,\cdots,s)$$
$$\times\sum_{r=1}^{s}J_{r,s+1}[f_1^{(0)}(s+1)]=0, \quad (s=1,2\cdots). \quad (30)$$

[This expression also contains the product of all the diagonal elements, for which we have already shown the truth of (30).] Now interpret $\delta_{y,0}$ as an even function of y, then its derivative is zero, and A_r does not act on the deltas in the product; this makes all deltas a multiplicative factor in (30). If all the y's are not equal to each other, this factor is zero, and the equation is satisfied; if all the y's have the same value, cancel the δ factors, and we are to satisfy the set (30) with the

[7] It is easy to see that the $(y_1-y_2)^{-2}$ type of singularity due to $\partial\phi(y_1-y_2)/\partial y_1$ does not affect this argument.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

stipulation $y_1=y_2=\cdots=y_s$. Now again, as before, the equation breaks up into a sum of terms, each of which vanishes, if (30) is satisfied for $s=1$. This is, however, satisfied if we require (28) to be true. Thus, in the limit $\epsilon=0$ we have solved the whole hierarchy of equations if we can solve (28) for $\epsilon=0$.

The Reduced Hierarchy and the Dirac Scheme

The problem is then solved if we find a solution of (28), (22), and (23) which minimizes the energy. Equation (28) is the equation of motion of the singlet density matrix; (22) expresses the two auxiliary conditions that the exclusion principle should be satisfied, and that the system should be in a pure state; (23) is a normalization condition. This set of equations are related to the set which formed the starting point of Dirac's investigations.[2] Dirac introduces a singlet density matrix, normalized to N, and stipulates that this density matrix should be a solution of the Hartree-Fock scheme with the subsidiary condition that the singlet density matrix should be idempotent, and that the energy be a minimum. Then he proceeds to solve this set of equations in the limit as Planck's constant tends to zero. Equation (28) is related to the equation of motion in the Hartree-Fock scheme, (22) to the idempotency condition, (23) to the normalization condition. However, in the further development there appears a disparity. Dirac solves in this limit the idempotency condition by an heuristic assumption on the form of the singlet density matrix; using the so resulting density matrix in the equations of motion he finds the TF solution, *amended by exchange terms, but with no other corrections*. This is astonishing for two reasons. First, according to our estimates in Sec. 3, the exchange terms are second-order corrections. Second, there is another second-order correction, the Weizsäcker correction. Thus, either no correction term should appear, or both. (Indeed, one of the more mysterious features of the Weizsäcker correction has always been that it somehow failed to appear in Dirac's approximation.) The inconsistency arises through the heuristic assumption used in solving the idempotency condition. This is a correct lower-order approximation; however, to this approximation the exchange terms in the equations of motion do also vanish [see (25)], the same way as the terms containing the Weizsäcker correction, and which indeed were tacitly put equal to zero by Dirac, when he equated the kinetic-energy term with its classical limit, $p^2/2m$.

For this reason in the next section we shall analyze the solution of our set of equations with somewhat greater care as usual, to demonstrate that in zeroth order the TF solutions are only obtained. To accomplish this aim we solve the equations in a manner which exhibits the similarities with the usual TF derivation as much as possible, instead of striking out along the path suggested by Dirac.

7. THE THOMAS-FERMI SOLUTION

We seek the solution of (28), (22), (23) which makes the energy a minimum. Let us start with (22) and (23). Fourier analyze $f_1^{(0)}$ in ζ:

$$f_1^{(0)}(y,\zeta)=(2\pi)^{-3}\int d^3k w(y,k)e^{ik\cdot\zeta}. \tag{31}$$

Equation (22) asserts that

$$w(y,k)^2=w(y,k), \tag{32}$$

or

$$w(y,k)=1, \quad \text{or} \quad 0. \tag{33}$$

The location of the boundary $h(y,k)=0$ separating the $w=1$ and $w=0$ domains (which can be multiply connected) is left undetermined by (22), while the normalization condition (23) merely requires that the volume of the $w=1$ region be $(2\pi)^3$. Observe that the determination of the boundary $h=0$ completes the solution of the problem. This is furnished by Eq. (28) and the minimal condition. Equation (28) written out in full reads:

$$\left\{\frac{\partial^2}{\partial y\partial\zeta}+\zeta\frac{\partial\phi(y)}{\partial y}\right\}f_1^{(0)}(y,\zeta)$$

$$-f_1^{(0)}(y,\zeta)\int d^3y_2\zeta\frac{\partial\phi(y-y_2)}{\partial y}f_1^{(0)}(y_2,0)=0. \tag{34}$$

Substitute (31) into (34) and observe that $w\zeta e^{ik\cdot\zeta}\equiv -iw(\partial e^{ik\cdot\zeta}/\partial k)\rightarrow i(\partial w/\partial k)e^{ik\cdot\zeta}$ by partial integration. We obtain:

$$k(\partial w/\partial y)-(\partial\Psi/\partial y)(\partial w/\partial k)=0, \tag{35a}$$

$$\Psi(y)=-\phi(y)+\int d^3y_2\phi(y-y_2)\int d^3k_2 w(y_2,k_2);$$

$$w(y,k)=1, \text{ or } 0. \tag{35b}$$

Equation (35) is the same as the classical Liouville equation for a one-particle distribution function $w(y,k)$, k being a dimensionless momentum variable, y a dimensionless position variable, and Ψ a dimensionless potential energy. The latter is a self-consistent potential since it depends on $\int d^3k w(y,k)$, which is proportional to the (yet unknown) density: thus (32) is the self-consistent field problem of Thomas and Fermi. Equation (35a) gives the equation of motion, (35b) the exclusion principle at T equal to zero. This is to be solved with the auxiliary condition that the total energy $\int d^3k \times\int d^3y(\frac{1}{2}k^2+\Psi)w$ be a minimum. What restrictions are placed by (35a) on the location of the boundary $h=0$? A straightforward substitution shows that $w(y,k)$ can only be a function of the combination $\frac{1}{2}k^2+\Psi(y)$. (In a stationary state the distribution function can only be a function of the constants of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

motion; the angular momentum being by hypothesis zero, only the energy $E=\frac{1}{2}k^2+\Psi$ remains.) Hence, the boundary $h=0$ must be a $E=$const surface, or several different energy surfaces pieced together.

To determine the boundary we must know (a) which energy surface, or surfaces, will be used, i.e., the determination of the constants in the $E=$const equation; (b) the potential energy Ψ.

We first answer (a). In the 6-dimensional (y,k) space (Fig. 3) one constructs the $E=$const surfaces; they will be spherically symmetrical in both the k three-space, and y three-space. The former is a consequence of E containing k^2 only; the latter is due to the fact that Ψ can only depend on $|y|$, since ϕ is a function of $|y|$, and so is $\int d^3k w(y,k)$ by symmetry. Were the angular momentum of the atom different from zero, a preferred plane would exist. (In the one-dimensional situation depicted in Fig. 3. This symmetry means that the $E=$const surfaces are symmetrical with respect to both the y and k axis.) By assumption only bound states should exist; thus $w(k,)y=W(E)$ must be zero if $E>0$; for $E<0$ $W(E)$ must be zero or one. In other words, we could have a W which consists of layers of 0 and 1, each layer bounded by $E=$const surfaces. (E.g., we could have on Fig. 2 $W=1$ between the $E=0$, $E=E'$ surfaces, $W=0$ between the $E=E'$ and $E=E''$ surfaces, etc.)

Now, it is possible to show that (a) the minimum energy obtains if there are no gaps, thus $W(E)=1$ if $E\leqslant E_c\leqslant 0$, where E_c is a critical energy; (b) the equations have no solution unless $E_c=0$. (The layered distributions correspond to excited states.) Thus

$$\begin{aligned} &W(E)=1 \quad E\leqslant 0,\\ &W(E)=0 \quad E>0,\\ &E=\tfrac{1}{2}k^2+\Psi(y). \end{aligned} \tag{36}$$

To answer (b) we proceed as follows: Take the Laplacian of Ψ with respect to y. We obtain

$$\Delta\Psi=4\pi\delta(y)-4\pi\int w(y,k)d^3k\,, \tag{37}$$

since $\Delta\phi=\Delta(1/|y|)=-4\pi\delta(y)$. The integral over w can be simply performed. Fix y; for each y, k can range over all values less than or equal to $k_m(y)$ defined by $\frac{1}{2}k_m^2+\Psi=0$; for all these values of k, w is equal to one; hence the integral is $(4\pi/3)k_m^3$. (The shaded area is Fig. 2.) Thus we get the two equations

$$\begin{aligned} &\Delta\Psi=-4\pi\delta(y)-4\pi(4\pi/3)k_m^3(y)\,,\\ &\tfrac{1}{2}k_m^2+\Psi=0. \end{aligned} \tag{38}$$

Equation (38) constitutes the standard set of Thomas-Fermi equations in dimensionless variables. The numerical factors are different. This is due to two trivial causes. First, we did not scale out 4π's and other numbers with our choice of L; second, as we mentioned, we omitted the factor 2 arising from the additional summation to be performed over the spin states. This then shows that in the limit $\epsilon\to 0$ the TF theory is a rigorous solution of the N-body problem.

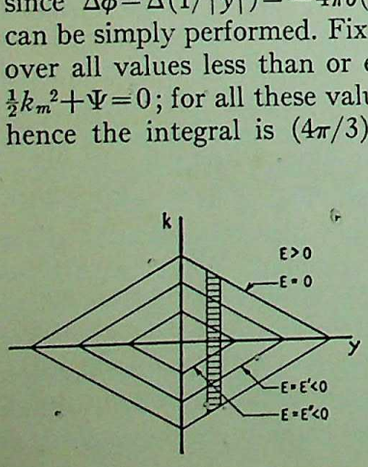


FIG. 3. Qualitative picture of (k,y) space, reduced μ space of electron, k-reduced momentum, y-reduced position; several constant-energy surfaces are shown.

8. THE INCONSISTENT NEXT APPROXIMATION

We have seen that the simple power-series expansion (15) does not exist. In spite of this it may be worthwhile to pretend temporarily its existence, in order to observe through what anomalies the equations inform us that we are attempting the impossible.

The pattern is clear. First, we have to assume a functional G which connects the f_s with f_1 to the next order in ϵ, still satisfying the exclusion principle. This, with the recursion relations gives a condition on f_1 which must be solved together with the equations of motion for the f_s. Let us assume that G is of the same form as F; thus, we extend to second order in ϵ the determinantal construction of the f_s from f_1. Then the recursion relations give again Eq. (21) as a condition, with $f_1(y,\zeta)=f_1^{(0)}(y,\zeta)+\epsilon^2 f_1^{(2)}(y,\zeta)$ in place of $f_1^{(0)}$, and with $y+\frac{1}{2}\epsilon\zeta$, $y-\frac{1}{2}\epsilon\zeta$ appearing in the arguments. Expanding it to second order in ϵ and equating the coefficient of ϵ^2 to zero, we get an inhomogeneous linear integral equation for $f^{(2)}$, where the homogeneous equation has only zero as its solution. Hence the inhomogeneous equation has, at most, one solution. This means that if a solution exists *we have found f_1 to second order* without ever investigating the equations of motion to second order. It is clear, however, that such a solution must be incompatible with the equations of motion for the following reason. We can modify the functional G in an infinite number of ways to give the same relation between f_1 and f_2, but different relations between f_1 and f_3, f_4, etc. This leaves the hypothetical solution unchanged but will modify the equations of motion. Thus, this specification of G is too restrictive. The solution, however, does not exist. If we formally solve the equation we find that the solution diverges. An inconsistent solution can be obtained, however. Assume that f_2 *is* a determinant, put

$$f_1(y,\zeta\,|\,\epsilon)=f_1^{(0)}(y,\zeta)+\epsilon^2 f_1^{(2)}(y,\zeta)$$

and expand it to second order in ϵ^2. Insert it in the equation of motion for f_1, and equate to zero separately the coefficient of ϵ^2. This way we obtain the previous zeroth-order equations for $f_1^{(0)}$, and an inhomogeneous equation for $f_1^{(2)}$. This equation is the one found by

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Kompaneets and Pavlovskii, and Baraff and Borowitz, containing the correction terms introduced earlier in a more heuristic way by Dirac (exchange correction) and by Weizsäcker (inhomogeneity correction). As we see, this result is then inconsistent in the sense that the exclusion principle is satisfied to zeroth order, while the equation of motion for f_1 is satisfied to second order; moreover it is left open how far the other equations of motion for f_s with $s>1$ are satisfied, since this involves a decision about the functional dependence of f_s on f_1.

9. CONJECTURES

As we see, the solution of the hierarchy of equations for the reduced density matrices involves three steps: the discovery of the proper small parameter; the establishment of a functional relation between the singlet density matrix and the higher density matrices; the invention of a calculational scheme to improve the approximations. We shall now comment on each step, bearing in mind what we have shown before. First, we observe the general difficulty that there is no general, all embracing postulate from which all three steps would necessarily follow. We first picked a small parameter such that in zeroth order we should recover the TF model; this of course does not make the small parameter unique. However, the parameter of smallness gives no clue as to the construction of the functional connecting the higher density matrices with the singlet density matrix. Moreover, this functional relation (chosen in such a way as to satisfy the exclusion principle) forces the small parameter already picked to appear in a nonanalytical manner. What we are missing is a postulate which connects the two separate assumptions about the small parameter and the functional relation. The basic feature of the TF model is that the energy is a functional of the density only. This unique feature should be incorporated from the beginning in some manner, as for example to seek those solutions of the hierarchy for which the singlet density matrix is a functional of the density (its diagonal elements) only, the pair density matrix a functional of the singlet density matrix, etc., and trying to determine the unknown functionals through the expansion in terms of a small parameter. Sidestepping this basic predicament, the present treatment suggests two basic improvements. We observe that the present small parameter essentially expresses the smallness of the off-diagonal elements of the singlet density matrix, but it does so in a very crude way. Think of the singlet density matrix in the x representation; our expansion says that only those elements are of importance which lie along the diagonal in a strip of *uniform* width ϵ. In reality this strip will be far from being uniform. As we go (in the x representation) along the diagonal, the diagonal elements $\rho(x,x)$ give the density as a function of the radial distance from the nucleus. The heuristic considerations in Sec. 3 show that the TF approximation is good around a distance L from the nucleus but poor everywhere else. Close to the nucleus the field is changing too rapidly; far away the density is insufficient. Thus, this strip is rather wide at both ends, becoming narrow in between. Thus, it is unwise to assign the small parameter ϵ uniformly in the (y,ζ) space; most likely there will be different asymptotic regions in this space with different correct assignment of smallness. One way to accomplish this is to introduce the unknown density distribution in the scale, and hence in the small parameter. Thus the small parameter must be specified implicitly. This already may make the nonanalyticity difficulties disappear. However, it seems to us that another remedy may also be required. In the present scheme, as it is usual, we insisted that the temperature of the atom is absolute zero. This may be too restrictive a stipulation. For, is it not rather preposterous to state, that you wish to determine the energy of the atom within an accuracy of, say, 20%, and at the same time to insist that you know its temperature exactly and not only to the same accuracy as you know the energy? It is likely that a coupled limiting process is required, in which the temperature is treated as an additional small parameter. Preliminary work suggests that this does indeed eliminate some of the difficulties, since at finite temperatures Eq. (22) acquires an additional term $kT\partial f/\partial\mu$ (μ being the chemical potential). Thus, we have essentially a "boundary layer" problem in μ around the Fermi surface, kT being the thickness of the boundary layer. It is a well-known fact of asymptotic analysis that one must *not* treat such cases by putting simply the boundary thickness equal to zero, since this obliterates the highest derivative in the equation.

10. SUMMARY AND CONCLUSIONS

We have shown then, that there exists a small parameter and a functional relation between the singlet and higher density matrices such that in the limit as this small parameter tends to zero, the TF solutions are exact solutions of the full N-body problem, for neutral atoms at absolute zero temperature. One surmises that the same would hold for ions as well, if under ionization we mean that a fraction of the total number of electrons are removed (and not a given number). The problem, however, involves the small parameter in question in a nonanalytical fashion and for this reason a simple power-series expansion cannot be used to go to the next approximation. If we are content to satisfy the exclusion principle to zeroth order, and the equations of motion of the singlet density matrix to next order, we obtain the equations found by Kompaneets and Pavlovskii, and Baraff and Borowitz, which contain the exchange and inhomogeneity correction terms. Finally, we suggest possible steps one could take to get rid of the difficulty generated by the nonanalyticities.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ACKNOWLEDGMENTS

It is a pleasure to thank all those who have helped in the completion of this work: part of it was supported by the National Science Foundation, part of it was done while I enjoyed the hospitality of the University of California, La Jolla. In particular, I owe my gratitude to Professor Joseph E. Mayer, who not only made my stay there possible but helped in a very essential way with important remarks and suggestions. Professor J. Lebowitz and Professor S. Borowitz kindly drew my attention to a number of references in the literature.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Theory of Core-Electron Contributions to Hyperfine Interactions*

G. D. GASPARI, WEI-MEI SHYU, AND T. P. DAS
Department of Physics, University of California, Riverside, California
(Received 5 August 1963; revised manuscript received 16 January 1964)

A convenient method is devised for the calculation of magnetic hyperfine constants in atoms, molecules, and metals taking into consideration the exchange interaction between the core electrons and the unpaired valence electrons. In this method, the core-electron wave functions are perturbed by the nuclear magnetic moment via the Fermi contact term, and the energy of the system is then calculated in the Hartree-Fock approximation using the perturbed core wave functions. The present method is closely related to the exchange perturbation method of Cohen, Goodings, and Heine. However, the former has the advantage of being more flexible in the sense that the same perturbed core-electron functions may be used for the ground and excited states of the atom and for metals without significant error. For lithium atom $1s^2 2s$ and $1s^2 2p$ states, we obtained values for the core contribution to the hyperfine constant a (in $a\mathbf{I}\cdot\mathbf{S}$) of 83.76 Mc/sec and -8.9 Mc/sec in good agreement with the earlier values of Cohen, Goodings, and Heine. We have applied this method to a calculation of the core-polarization correction to the Knight shift in lithium metal using recent wave functions of Kohn and Callaway. The core-polarization corrections produced by the s and p parts of the conduction-electron wave function are nearly equal but opposite in sign, while that produced by the d part is an order of magnitude smaller. This results in a net correction of about -5.3% of the direct contribution to the Knight shift from the conduction electrons.

I. INTRODUCTION

A NUMBER of recent detailed and careful investigations[1–10] of the magnetic hyperfine interaction in atoms and paramagnetic ions have clearly demonstrated the important role of exchange polarization of the core electrons in contributing to the magnetic hyperfine interaction constant. In all these investigations, the method that has been employed has come to be known generally as the unrestricted Hartree-Fock (UHF) method. In keeping with the recent attempt of standardization of nomenclature,[11] we shall call the unrestricted Hartree-Fock method the spin polarized method (SP). Similarly, the projected unrestricted Hartree-Fock method will be denoted as the projected spin polarized method (PSP).[12] Some of the investigators[2–4,6,9] have handled the SP method self-consistently, while others have used a perturbation approach. To avoid confusion we shall refer to the method of treating the exchange potential as a perturbation, the exchange perturbation method (EP). In recent papers, Nesbet,[7] Marshall,[8] and Heine[13] have discussed possible errors that can occur in the results of calculation by the SP method, because the many-electron wave function used in the SP method is not an eigenfunction of $\mathbf{S}^2$, where $\mathbf{S}$ is the total spin of the atom. From the investigations of these authors, one arrives at the conclusion, that for paramagnetic ions and atoms, this limitation of the SP method is not a serious source of error; but in extending the SP method to metals and molecules, one has to be more careful about the influence of this source of error.

In this paper we are interested in a perturbation method which has the advantage of flexibility over the EP method while not sacrificing accuracy. The accuracy

---

* Supported by the National Science Foundation.
[1] R. M. Sternheimer, Phys. Rev. **86**, 316 (1952).
[2] V. Heine, Phys. Rev. **107**, 1002 (1957).
[3] J. H. Wood and G. W. Pratt, Phys. Rev. **107**, 995 (1957).
[4] D. A. Goodings, Phys. Rev. **123**, 1706 (1961).
[5] M. H. Cohen, D. A. Goodings, and V. Heine, Proc. Phys. Soc. (London) **73**, 811 (1959).
[6] R. E. Watson and A. J. Freeman, Phys. Rev. **120**, 1125, 1134 (1960).
[7] R. K. Nesbet, Proc. Roy. Soc. (London) **A230**, 312 (1955); Phys. Rev. **118**, 681 (1960).
[8] W. Marshall, Proc. Phys. Soc. (London) **78**, 113 (1961); see also N. Bessis, H. Lefebvre-Brion, and C. M. Moser, Phys. Rev. **124**, 1124 (1961).
[9] L. M. Sachs, Phys. Rev. **117**, 1504 (1960).
[10] K. F. Berggren and R. F. Wood, Phys. Rev. **130**, 198 (1963).
[11] N. Bessis, H. Lefebvre-Brion, and C. M. Moser, Rev. Mod. Phys. **35**, 548 (1963).
[12] The PSP method should be distinguished from the SPP method. In the SPP method one applies the projection operator after an SP calculation is performed to obtain an eigenfunction of $\mathbf{S}^2$.
[13] V. Heine (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the method can be easily compared with other perturbation calculations, which we shall show to be exactly equivalent in principle to our method. Specifically, what one does is to perturb the core electrons with the nuclear magnetic moment (μ) via the Fermi contact interaction; and with these perturbed functions, one then computes the energy to first order in μ, which is directly related to the hyperfine coupling constant. This method, which we shall call the moment perturbation method (MP), differs from the EP approach only in the order of application of the perturbations, since there the core electrons are perturbed by the exchange potential rather than the nuclear magnetic moment. The flexibility of the method arises by letting the nucleus perturb the core functions and approximating (which we demonstrate by our lithium atom calculations to be a good approximation) these perturbed functions as being independent of the configuration of the outer valence electrons. Thus, we assume that the interaction between the magnetic moment of the nucleus and the core electrons of the atom is not greatly altered when one goes from the atomic ground state to an excited state, or to molecular and metallic states. This has the advantage of enabling one to calculate exchange-polarization effects when the system is in a variety of environments without having to solve separate differential equations in each case. In order to demonstrate the practicality and flexibility of the MP method, we have calculated the hyperfine interaction constants for atomic lithium in the $1s^22s$ and $1s^22p$ configurations, and compared them with the results of earlier calculations. We then apply the MP method to calculate the contribution of exchange polarization to Knight shift in metallic lithium. In all these calculations, the perturbed core wave functions used were those calculated for atomic lithium $1s^22p$ configuration. This procedure[14] is analogous to similar cross perturbation methods that have been used for the quadrupole antishielding factors for ions,[15] and for the calculation of nuclear magnetic shielding coefficients in molecules.[16]

In Sec. II we describe the details of our method and its correspondence with the EP method. In Sec. III the method is applied to lithium atom $1s^22s$ and $1s^22p$ states, and the results are compared with earlier EP calculations. In Sec. IV we calculate the core-polarization correction to the Knight shift in lithium metal using the recent Wigner-Seitz wave functions obtained by Kohn and Callaway.[17] In Sec. V we shall discuss the limitations of the MP method and the problem of self-consistency.

THEORY OF THE MP METHOD

The general procedure followed here is similar to that employed in some other perturbation problems. However, for the sake of completeness, we have presented here an adaptation pertinent to our problem at hand. Let us consider a general system whose unperturbed Hamiltonian is $\mathcal{H}_0$ and whose ground-state wave function is Ψ_0.

$$\mathcal{H}_0\Psi_0=E_0\Psi_0. \tag{1}$$

Let the system be perturbed by two general first-order perturbing forces described by Hamiltonians $\mathcal{H}_E$ and $\mathcal{H}_N$. If $\delta\Psi_E$ and $\delta\Psi_N$ represent the first-order changes in the wave function of the system due to $\mathcal{H}_E$ and $\mathcal{H}_N$, respectively, then the energy of the system correct to the second order is

$$E=\frac{\langle\Psi_0+\delta\Psi_E+\delta\Psi_N|\mathcal{H}_0+\mathcal{H}_E+\mathcal{H}_N|\Psi_0+\delta\Psi_E+\delta\Psi_N\rangle}{\langle\Psi_0+\delta\Psi_E+\delta\Psi_N|\Psi_0+\delta\Psi_E+\delta\Psi_N\rangle}. \tag{2}$$

On simplication, after making a binomial expansion of the denomintaor, one gets

$$\begin{aligned}E=\{&E_0+E_e+E_N+2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_E\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_E\rangle\\&+2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_N\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_N\rangle+\langle\delta\Psi_E|\mathcal{H}|\delta\Psi_E\rangle\\&+\langle\delta\Psi_N|\mathcal{H}|\delta\Psi_N\rangle+2\langle\delta\Psi_E|\mathcal{H}_0|\delta\Psi_N\rangle\}\\&-\{\langle\delta\Psi_E|\delta\Psi_E\rangle-\langle\delta\Psi_N|\delta\Psi_N\rangle-2\langle\delta\Psi_E|\delta\Psi_N\rangle\}.\end{aligned} \tag{3}$$

After omitting terms higher than second order, Eq. (3) becomes

$$\begin{aligned}E=&E_0+E_e+E_N+\langle\delta\Psi_N|\mathcal{H}_0-E_0|\delta\Psi_N\rangle\\&+\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_E\rangle+2\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_N\rangle\\&+2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_E\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_E\rangle\\&+2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_N\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_N\rangle,\end{aligned} \tag{4}$$

where we have assumed that the perturbed wave function is normalized to first order. That is,

$$\langle\Psi_0|\delta\Psi_E\rangle=\langle\Psi_0|\delta\Psi_N\rangle=0, \tag{5}$$

and the first-order changes in energy are given by

$$\begin{aligned}E_e&=\langle\Psi_0|\mathcal{H}_E|\Psi_0\rangle,\\E_N&=\langle\Psi_0|\mathcal{H}_N|\Psi_0\rangle.\end{aligned} \tag{6}$$

The second-order change in energy is given by

$$\begin{aligned}E^{(2)}=&2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_E\rangle+\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_E\rangle\\&+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_N\rangle+\langle\delta\Psi_N|\mathcal{H}_0-E_0|\delta\Psi_N\rangle\\&+2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_N\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_E\rangle\\&+2\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_N\rangle.\end{aligned} \tag{7}$$

[14] While this work was in progress, we received a preprint of a paper by Heine (Ref. 13) in which he mentioned that E. Šimánek had independently proposed a similar method to Heine to avoid some of the inaccuracies of the SP method that were discussed in Ref. 13.

[15] R. M. Sternheimer, Phys. Rev. 84, 244 (1951); R. M. Sternheimer and H. M. Foley, *ibid.* 102, 731 (1956); T. P. Das and R. Bersohn, *ibid.* 102, 733 (1956).

[16] T. P. Das and R. Bersohn, Phys. Rev. 115, 897 (1959).

[17] W. Kohn and J. Callaway, Phys. Rev. 127, 1913 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The terms in the second-order change in energy can be regrouped[18] as

$$E_e^{(2)}=2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_E\rangle+\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_E\rangle,$$
$$E_N^{(2)}=2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_N\rangle+\langle\delta\Psi_N|\mathcal{H}_0-E_0|\delta\Psi_N\rangle,$$
$$E_{eN}^{(2)}=2\langle\Psi_0|\mathcal{H}_E|\delta\Psi_N\rangle+2\langle\Psi_0|\mathcal{H}_N|\delta\Psi_E\rangle+2\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_N\rangle. \quad (8)$$

The quantity $E_e^{(2)}$ is the second-order change in energy due to $\mathcal{H}_E$ alone; $E_N^{(2)}$ is the second-order change in energy due to $\mathcal{H}_N$ alone; and $E_{eN}^{(2)}$ is the second-order change in energy due to the interaction of $\mathcal{H}_E$ and $\mathcal{H}_N$.

The first-order equations for $\delta\Psi_E$ and $\delta\Psi_N$ are

$$(\mathcal{H}_0-E_0)\delta\Psi_E=-(\mathcal{H}_E-E_e)\Psi_0, \quad (9)$$

$$(\mathcal{H}_0-E_0)\delta\Psi_N=-(\mathcal{H}_N-E_N)\Psi_0, \quad (10)$$

which may also be derived by minimizing $E_e^{(2)}$ and $E_N^{(2)}$ with respect to $\delta\Psi_E$ and $\delta\Psi_N$, respectively. From Eqs. (9) and (10) it is easily seen that

$$\langle\delta\Psi_N|\mathcal{H}_E|\Psi_0\rangle=\langle\delta\Psi_E|\mathcal{H}_N|\Psi_0\rangle=-\langle\delta\Psi_E|\mathcal{H}_0-E_0|\delta\Psi_N\rangle. \quad (11)$$

Hence, using Eq. (11), $E_{eN}^{(2)}$ reduces to

$$E_{eN}^{(2)}=2\langle\delta\Psi_E|\mathcal{H}_N|\Psi_0\rangle, \quad (12)$$

$$E_{eN}^{(2)}=2\langle\delta\Psi_N|\mathcal{H}_E|\Psi_0\rangle. \quad (13)$$

One could now apply this result to our problem of core polarization. For the perturbing Hamiltonian $\mathcal{H}_E$, we take the exchange potential at the position of the core electron due to the unpaired valence electron and for $\mathcal{H}_N$, the Fermi contact interaction due to the nuclear magnetic moment.

$$\mathcal{H}_E=-\frac{\psi_{\text{val}}(1)}{\psi_{\text{core}}(1)}\int\psi_{\text{val}}(2)\frac{e^2}{r_{12}}\psi_{\text{core}}(2)d\tau_2, \quad (14)$$

$$\mathcal{H}_N=\frac{16\pi}{3}\gamma_e\gamma_N\hbar^2\mathbf{I}\cdot\mathbf{S}\delta(\mathbf{r}). \quad (15)$$

In Eq. (15), γ_e is the magnetogyric ratio ($e/2$ Mc/sec) of the electron, γ_N is the magnetogyric ratio of the nucleus, and $\mathbf{I}$ and $\mathbf{S}$ are the respective spins of the nucleus and electron. Equation (14) requires some explanation[5]:

In the presence of an unpaired valence electron, with spin "up," Eq. (14) represents the difference in exchange potential seen by the up and down core electrons. If the restricted Hartree-Fock Hamiltonian for the core states were used as $\mathcal{H}_0$ in Eq. (9), we should consider two first-order equations of the type of Eq. (9), one for the up core state and one for the down. The equation for the up state would involve $+\mathcal{H}_E/2$ on the right and that for the down state $-\mathcal{H}_E/2$. For those effects which involve $\mathcal{H}_E$ to first order, one can consider the down state to remain unperturbed and the up state to be perturbed by $\mathcal{H}_E$ as given in Eq. (14).

Besides $E_{eN}^{(2)}$, the various higher order terms that could contribute to the hyperfine interaction are $\mathcal{H}_E^2\mathcal{H}_N$, $\mathcal{H}_E^3\mathcal{H}_N$, $\mathcal{H}_E^4\mathcal{H}_N$, $\cdots$. These terms are expected to be negligible in heavy atoms where $\mathcal{H}_E$ is a small perturbation compared to $\mathcal{H}_0$. For light atoms there may be a non-negligible contribution from such terms. We have not investigated its effect in lithium atom since our major interest is in Knight shift calculations. It will be shown in Sec. IV that these higher order terms in $\mathcal{H}_E$ do not have to be considered in core-polarization contributions to Knight shift. Higher order terms like $\mathcal{H}_N^2\mathcal{H}_E$, $\mathcal{H}_N^3\mathcal{H}_E$, $\cdots$, lead to terms in the energy non-linear in nuclear spin and are of no interest to us.

For metal ψ_{val} becomes ψ_{cond}, the wave function for a conduction electron which extends over the entire crystal and therefore, involves more than one center. The perturbation equation for $\delta\Psi_E$ [Eq. (9)] cannot, therefore, be separated into radial and angular parts. Cohen, Goodings, and Heine[5] attempted to meet this limitation by expressing ψ_{cond} around the nucleus in question as a linear combination of spherical harmonics and used the linear independence of spherical harmonics of different orders to separate Eq. (9) into a number of independent differential equations. This procedure is rather cumbersome because one now has to solve a set of differential equations in place of Eq. (9). In the MP procedure, one considers the perturbations in the reverse order. This is justified because of the two alternative expressions (12) and (13) for $E_{eN}^{(2)}$. Instead of first considering the perturbation $\delta\Psi_E$ due to $\mathcal{H}_E$, let us consider the perturbation of the wave function $\delta\Psi_N$ due to $\mathcal{H}_N$. Thus, we can determine $\delta\Psi_N$ by solving Eq. (10) and obtain $E_{eN}^{(2)}$ from Eq. (13). Since $\mathcal{H}_N$ is a localized perturbation by a point source and ψ_0 refers to ψ_{core} (the wave function of the core electrons) which is, to a very good approximation, localized and centrosymmetric, Eq. (10) can be separated into radial and angular parts. It should be noted that when the wave function for the conduction electron is expanded as a sum of spherical harmonics, the MP method obviates the necessity of having to solve a set of differential equations; but instead, one has to evaluate a set of integrals arising out of Eq. (13). In principle, the two procedures for calculating $E_{eN}^{(2)}$ should lead to identical results; but in practice, this will not be exactly true. In the process of obtaining numerical solutions of the differential Eqs. (9) and (10), one weights various regions of space differently because of the different natures of the two perturbations $\mathcal{H}_E$ and $\mathcal{H}_N$. This point will become clear in Sec. III, where we compare our results with those of Cohen, Goodings, and Heine's EP calculations on lithium atom. We shall also show in Sec. III that for a model, where analytic solutions are obtainable with both EP and MP methods, the two methods do, in fact, lead to identical results.

[18] T. P. Das and A. Mukherjee, J. Chem. Phys. 33, 1808 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

With these general remarks, we now proceed to the detailed forms of the equations that one has to solve for the MP method. Since, in a general case, one may have more than one core state to perturb, a formulation such as Dalgarno's for the perturbation of many-electron systems[19] is helpful. We will follow Dalgarno's method II (i.e., non-self-consistent perturbation) and obtain the first-order perturbation equations for $\delta\psi_N$ by minimizing the second-order energy [Eq. (59) in Dalgarno's paper] with respect to $\delta\psi_N$.

The first-order perturbation equations are then given by

$$(\mathcal{H}_i-\epsilon_i)\delta\psi_{i,N}-\sum_j(\epsilon_j-\epsilon_i)\langle\psi_j|\delta\psi_{i,N}\rangle\psi_j = -\mathcal{H}_N\psi_i+\sum_j\langle\psi_j|\mathcal{H}_N|\psi_i\rangle\psi_j, \quad (16)$$

where $\mathcal{H}_i$ is the one-electron Hamiltonian, and ψ_i, ψ_j are the one-electron wave functions in the Hartree-Fock approximation, and ϵ_i and ϵ_j are the corresponding eigenvalues. The first-order change in the wave function for the ith one-electron state due to $\mathcal{H}_N$ is $\delta\psi_{i,N}$. For a core ns state, Eq. (16) for $\delta\psi_{ns,N}$ takes the form

$$(\mathcal{H}_{ns}-\epsilon_{ns})\delta\psi_{ns,N}=\sum_{n's}(\epsilon_{n's}-\epsilon_{ns})\langle\psi_{n's}|\delta\psi_{ns,N}\rangle\psi_{n's} -\mathcal{H}_N\psi_{ns}+\sum_{n's}\langle\psi_{n's}|\mathcal{H}_N|\psi_{ns}\rangle\psi_{n's}. \quad (17)$$

One has to solve Eq. (17) for all occupied core ns states, since the s states are the only ones which do not have nodes at the origin and, hence, can contribute to the hyperfine interaction through the Fermi contact term. The summation terms on the right-hand side extend over all the core s states.

The zero-order one-electron wave function ψ_{ns} satisfies the equation

$$(\mathcal{H}_{ns}-\epsilon_{ns})\psi_{ns}=0. \quad (18)$$

That is, in atomic units

$$(-\nabla^2+V_{ns}-\epsilon_{ns})\psi_{ns}=0, \quad (19)$$

where V_{ns} is the one-electron potential seen by an electron in the ns state in the Hartree-Fock approximation. Therefore,

$$(V_{ns}-\epsilon_{ns})=\nabla^2\psi_{ns}/\psi_{ns}. \quad (20)$$

Substituting this in Eq. (17), we get

$$\left(-\nabla^2+\frac{\nabla^2\psi_{ns}}{\psi_{ns}}\right)\delta\psi_{ns,N} =\sum_{n's}(\epsilon_{n's}-\epsilon_{ns})\langle\psi_{n's}|\delta\psi_{ns,N}\rangle\psi_{n's} -\mathcal{H}_N\psi_{ns}+\sum_{n's}\langle\psi_{n's}|\mathcal{H}_N|\psi_{ns}\rangle\psi_{n's}. \quad (21)$$

[19] A. Dalgarno, Proc. Roy. Soc. (London) **A251**, 282 (1959).

If we now introduce the new function $\phi_{ns,N}$, defined by

$$\delta\psi_{ns,N}=-\frac{A}{4\pi}\frac{\psi_{ns}}{r}+\frac{\phi_{ns,N}}{r}Y_0^0(\theta,\phi), \quad (22)$$

where

$$A=\frac{16\pi}{3}\gamma_e\gamma_N\hbar^2 I_z\frac{m}{a_0\hbar^2},$$

and remembering that $\nabla^2(1/r)=-4\pi\delta(\mathbf{r})$, Eq. (21) reduces to

$$-\frac{d^2\phi_{ns,N}}{dr^2}+\frac{A}{2\pi r^3}u_{ns}-\frac{A}{2\pi r^2}\frac{du_{ns}}{dr}+\frac{\phi_{ns,N}}{u_{ns}}\frac{d^2u_{ns}}{dr^2} =\sum_{n's}(\epsilon_{n's}-\epsilon_{ns})\{\langle u_{n's}|\phi_{ns,N}\rangle u_{n's}\} +A\sum_{n's}\psi_{n's}(0)\psi_{ns}(0)u_{n's}, \quad (23)$$

where we have used the notation

$$\psi_{ns}=[u_{ns}(r)/r]Y_0^0(\theta,\phi).$$

Because of the first term on the right-hand side, Eq. (23) is an integrodifferential equation and must be solved self-consistently. This term represents the influence of Pauli principle on the perturbed core states.

After determining $\delta\psi_{ns,N}$ by solving Eq. (23), one can get the second-order energy $E_{eN}^{(2)}$ using Eq. (13). Thus, substituting for the determinantal function $\delta\Psi_N$ in terms of the one-electron functions $\delta\psi_{ns,N}$ and using Eq. (14) for $\mathcal{H}_E$, we get

$$E_{eN}^{(2)}=2\sum_n\left\{\int\psi_{\rm val}(1)\delta\psi_{ns,N}(1)\psi_{\rm val}(2)\psi_{ns}(2) \times\frac{1}{r_{12}}d\tau_1d\tau_2-\sum_{n'}\int\psi_{\rm val}(1)\psi_{ns}(1)\psi_{\rm val}(2)\psi_{n's}(2) \times\frac{1}{r_{12}}d\tau_1d\tau_2\int\psi_{n's}\delta\psi_{ns,N}d\tau\right\}. \quad (24)$$

From this equation the correction a_c to the hyperfine constant a from core polarization is obtained by the equation

$$a_c=E_{eN}^{(2)}/IJ2\pi\hbar. \quad (25)$$

In solids and molecules, **J** is equal to the total spin **S**; whereas, for atoms, **J** is the total angular momentum of the valence electrons.

III. RESULTS FOR LITHIUM ATOM $1s^22p$ AND $1s^22s$ STATES

For the lithium atom $1s^22p$ ($J=\frac{3}{2}$) state Eq. (23) for $\phi_{1s,N}$ takes the form

$$-\frac{d^2\phi_{1s,N}}{dr^2}+\frac{A}{2\pi r^3}u_{1s}-\frac{A}{2\pi r^2}\frac{du_{1s}}{dr}+\frac{\phi_{1s,N}}{u_{1s}}\frac{d^2u_{1s}}{dr^2} =A|\psi_{1s}(0)|^2u_{1s}. \quad (26)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To illustrate the nature of the solution of Eq. (26), we first assume a hydrogenic form for u_{1s} for which an exact solution of Eq. (23) can be obtained. The Coulson-Duncanson functions[20] represent such a choice.

$$\psi_{1s}=\left(\frac{z^3}{\pi}\right)^{1/2}e^{-zr},\quad \psi_{2p_0}=\left(\frac{\xi^5}{\pi}\right)^{1/2}re^{-\xi r}\cos\theta \tag{27}$$

and

$$z=2.69,\quad \xi=0.525.$$

The solution of Eq. (21) is then found to be[21]

$$\phi_{1s,N}=\frac{A}{4\pi}(2z\ln r+2z^2r)u_{1s}(r). \tag{28}$$

To find a solution $\delta\psi_{1s,N}$ which satisfies Eq. (5), we use the Schmidt orthogonalization procedure and obtain from Eqs. (22) and (28)

$$\delta\psi_{1s,N}=\frac{A}{4\pi}[F_{1s}-\langle\psi_{1s}|F_{1s}|\psi_{1s}\rangle]\psi_{1s},$$

where (29)

$$F_{1s}=-\frac{1}{r}+2z\ln r+2z^2r.$$

From Eq. (14), the exchange perturbation $\mathcal{K}_E$ due to the $2p_0$ electron is given by

$$\mathcal{K}_E=-\frac{u_{2p}(r)}{3u_{1s}(r)}\left[\frac{1}{r^2}\int_0^r u_{2p}(s)u_{1s}(s)sds+r\int_r^\infty u_{2p}(s)u_{1s}(s)\frac{1}{s^2}ds\right]. \tag{30}$$

Hence, from Eqs. (24) and (30) it follows that

$$E_{eN}^{(2)}=-\tfrac{2}{3}B\left\{\int_0^\infty u_{2p}(r)\delta u_{1s,N}(r)dr\times\left[\frac{1}{r^2}\int_0^r u_{1s}(s)u_{2p}(s)sds+r\int_r^\infty u_{1s}(s)u_{2p}(s)\frac{1}{s^2}ds\right]\right\}, \tag{31}$$

where

$$B=\frac{A}{4\pi}\times\frac{e^2}{a_0}=\frac{4}{3}\frac{\gamma_N\gamma_e\hbar^2I_z}{a_0^3},$$

and

$$\delta u_{1s,N}=\frac{4\pi}{A}r\delta\psi_{1s,N}. \tag{32}$$

For the Coulson-Duncanson functions using Eqs. (27), (28), and (31), and performing the necessary integrations, we get

$$E_{eN}^{(2)}=-\frac{32}{9}B\xi^5z^3\left[\frac{147z^2}{4(z+\xi)^8}-\frac{15}{8(z+\xi)^6}-\frac{273z}{8(z+\xi)^7}-\frac{21z\ln(z+\xi)}{2(z+\xi)^7}+\frac{48z\ln 2}{(z+\xi)^7}+\frac{21z\ln z}{2(z+\xi)^7}\right]. \tag{33}$$

Substituting the values of z and ξ, one gets

$$E_{eN}^{(2)}=-0.0540B. \tag{34}$$

From Eqs. (25) and (32), remembering $I=J=\frac{3}{2}$, it then follows that

$$a_c=\frac{4\mu_N\mu_0E_{eN}^{(2)}}{3a_0^3IJ(2\pi\hbar)B}=-5.24\ \text{Mc/sec}, \tag{35}$$

where $\mu_N=\gamma_NI\hbar$ is the magnetic moment of the Li^7 nucleus and μ_0 is the Bohr magneton.

Since $u_{1s}(r)$ involves a single exponential for the Coulson-Duncanson function, one can obtain an exact analytic solution for the perturbation equation (16) for the EP method, with $\delta\psi_{iN}$ and $\mathcal{K}_N$ replaced by $\delta\psi_{iE}$ and $\mathcal{K}_E$, respectively.

$$\delta\psi_{1s,E}=-\frac{8}{9}\left(\frac{z^3}{\pi}\right)^{1/2}\frac{\xi^5e^{-zr}}{(z+\xi)^7}\left\{\frac{3(z+\xi)^4r^2e^{-2\xi r}}{4\xi}+\left[\frac{3(z+\xi)^4}{4\xi^2}+\frac{9(z+\xi)^3}{2\xi}\right]re^{-2\xi r}+\left[\frac{3(z+\xi)^4}{8\xi^3}+\frac{9(z+\xi)^3}{4\xi^2}+\frac{51(z+\xi)^2}{4\xi}\right]e^{-2\xi r}+\frac{75e^{-2\xi r}}{4r}-\frac{24e^{(z-\xi)r}}{r}-\frac{75z}{2}\int\frac{e^{-2\xi r}}{r}dr+48z\int\frac{e^{(z-\xi)r}}{r}dr\right\}-\frac{28}{3}\left(\frac{z^3}{\pi}\right)^{1/2}\frac{z^2\xi^5e^{-zr}}{(z+\xi)^7}\left(r+\frac{1}{z}\ln r-\frac{1}{2z^2r}\right). \tag{36}$$

Substituting $\delta\psi_{1s,E}$ from Eq. (36) into Eq. (12), one obtains, after some manipulation, exactly the same expression for $E_{eN}^{(2)}$ as in Eq. (33).

The value of a_c in Eq. (35) is about one-half of Goodings'[4] calculated value (−10.7 Mc/sec) by the SP method and is in somewhat better agreement with Cohen, Goodings, and Heine's[5] value (−8.5 Mc/sec) by the EP method. However, one does not expect very good agreement with these earlier results because the single-exponential hydrogen functions are not good approximations to the Hartree-Fock wave functions.[22]

To get a result by the MP method which could be properly compared in terms of accuracy with Cohen,

[20] C. Coulson and W. E. Duncanson, Proc. Roy. Soc. (Edinburgh) 62, 37 (1944).

[21] C. Schwartz, Ann. Phys. (N. Y.) 6, 156 (1959).

[22] Per-Olov Löwdin and K. Appel, Phys. Rev. 103, 1746 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

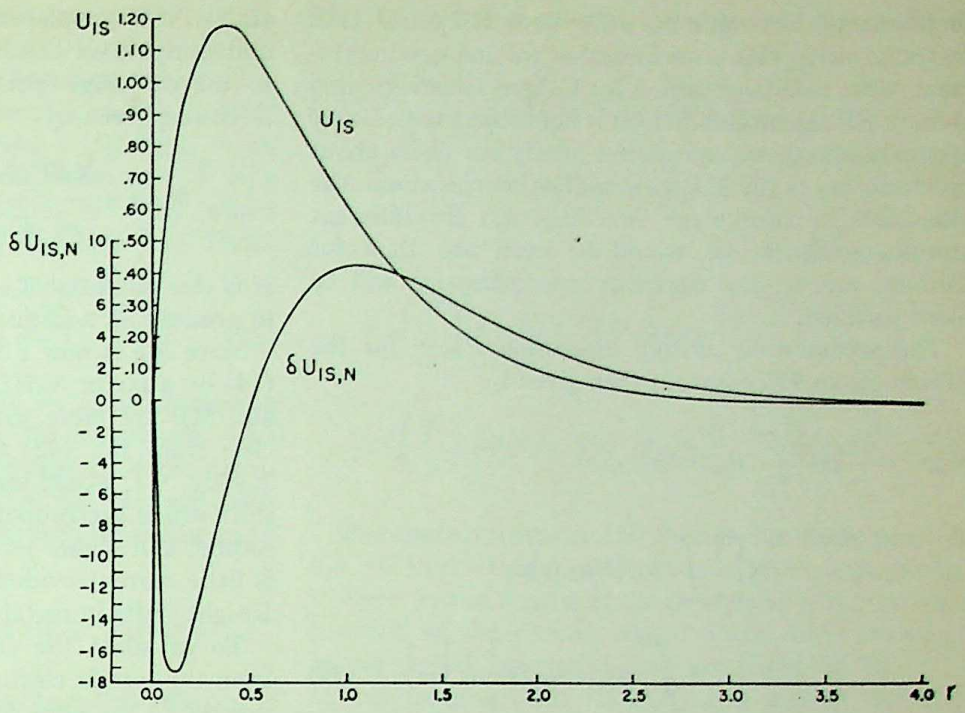


FIG. 1. Plot of unperturbed (u_{1s}) wave function and first-order perturbation $\delta u_{1s,N}$ as function of r for lithium atom $1s^22p$ configuration.

Goodings, and Heine's EP calculation and Goodings' SP calculation, we next solved Eq. (23) using Hartree-Fock wave functions. The solution has now to be obtained by numerical integration of Eq. (23). In Fig. 1, the function $\delta u_{1s,N}$ [defined in Eq. (32)] is plotted, together with Goodings' Hartree-Fock wave function[20] u_{1s}, which was employed in our calculations. The function $\delta u_{1s,N}$ resembles a $2s$ type function, in that, it has a node, but $\delta\psi_{1s,N}=(A/4\pi)(\delta u_{1s,N}/r)$ goes to negative infinity at the origin.

Using Eq. (31), together with Goodings' HF wave functions u_{1s}, u_{2p}, and $\delta u_{1s,N}$ of Fig. 1, it follows that

$$E_{cN}^{(2)}=-\tfrac{2}{3}B(0.137). \tag{37}$$

From Eq. (25) this leads to

$$a_c=-8.9 \text{ Mc/sec}. \tag{38}$$

This value compares very favorably with Cohen, Goodings, and Heine's EP value (Table I). This is satisfying in light of our Eqs. (12) and (13) and our earlier demonstration that EP and MP methods lead to identical results whenever exact solutions of Eqs. (9) and (10) are obtainable.

Both ours and Cohen, Goodings, and Heine's values for a_c are in slight disagreement with Goodings' value (Table I) obtained by the SP method. However, no measured value of a_c is available, so it is rather difficult to assess the importance of this small disagreement between the values obtained by MP (and EP) and SP methods.

The feasibility and accuracy of the MP method has now been demonstrated by its favorable comparison with the results of the EP method for the lithium atom $1s^22p$ state. We would next like to show the flexibility of the MP method by a consideration of core polarization effects for the $1s^22s$ state of lithium atom. The perturbation equation to be solved to get $\delta\psi_{1s,N}$ (or $\phi_{1s,N}$) in this case is exactly the same as in Eq. (26) except that u_{1s} may be different from its value for the $1s^22p$ state. Available tables[23] of u_{1s} for $1s^22s$ and $1s^22p$ states show that there is only a very slight difference between the values of u_{1s} in the two cases. One could

TABLE I. List of contributions to hyperfine constants in Mc/sec for lithium atom $1s^22s$ and $1s^22p$ states.

	$1s^22s$			$1s^22p$		
	Core polarization	Direct $2s$ contribution	Total	Core polarization	Direct $2p$ contribution	Total
Goodings	106	284	390	−10.7	0	−10.7
Cohen, Goodings, and Heine[a]	86.74	286.26	373.0	−8.5	0	−8.5
This paper	83.76	284	367.76	−8.9	0	−8.9
Experimental			401.786			

[a] The direct $2s$ contribution is taken from Table I of Ref. 5.

[23] D. A. Goodings, Ph.D. thesis, Cambridge University, Cambridge, England, 1961 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To illustrate the nature of the solution of Eq. (26), we first assume a hydrogenic form for u_{1s} for which an exact solution of Eq. (23) can be obtained. The Coulson-Duncanson functions[20] represent such a choice.

$$\psi_{1s}=\left(\frac{z^3}{\pi}\right)^{1/2}e^{-zr},\quad \psi_{2p_0}=\left(\frac{\xi^5}{\pi}\right)^{1/2}re^{-\xi r}\cos\theta \qquad (27)$$

and

$$z=2.69,\quad \xi=0.525.$$

The solution of Eq. (21) is then found to be[21]

$$\phi_{1s,N}=\frac{A}{4\pi}(2z\ln r+2z^2r)u_{1s}(r). \qquad (28)$$

To find a solution $\delta\psi_{1s,N}$ which satisfies Eq. (5), we use the Schmidt orthogonalization procedure and obtain from Eqs. (22) and (28)

$$\delta\psi_{1s,N}=\frac{A}{4\pi}[F_{1s}-\langle\psi_{1s}|F_{1s}|\psi_{1s}\rangle]\psi_{1s},$$

where (29)

$$F_{1s}=-\frac{1}{r}+2z\ln r+2z^2r.$$

From Eq. (14), the exchange perturbation $\mathcal{K}_E$ due to the $2p_0$ electron is given by

$$\mathcal{K}_E=-\frac{u_{2p}(r)}{3u_{1s}(r)}\left[\frac{1}{r^2}\int_0^r u_{2p}(s)u_{1s}(s)sds+r\int_r^\infty u_{2p}(s)u_{1s}(s)\frac{1}{s^2}ds\right]. \qquad (30)$$

Hence, from Eqs. (24) and (30) it follows that

$$E_{eN}^{(2)}=-\tfrac{2}{3}B\left\{\int_0^\infty u_{2p}(r)\delta u_{1s,N}(r)dr\times\left[\frac{1}{r^2}\int_0^r u_{1s}(s)u_{2p}(s)sds+r\int_r^\infty u_{1s}(s)u_{2p}(s)\frac{1}{s^2}ds\right]\right\}, \qquad (31)$$

where

$$B=\frac{A}{4\pi}\times\frac{e^2}{a_0}=\frac{4}{3}\frac{\gamma_N\gamma_e\hbar^2I_z}{a_0^3},$$

and

$$\delta u_{1s,N}=\frac{4\pi}{A}r\delta\psi_{1s,N}. \qquad (32)$$

For the Coulson-Duncanson functions using Eqs. (27), (28), and (31), and performing the necessary integrations, we get

$$E_{eN}^{(2)}=-\frac{32}{9}B\xi^5z^3\left[\frac{147z^2}{4(z+\xi)^8}-\frac{15}{8(z+\xi)^6}-\frac{273z}{8(z+\xi)^7}-\frac{21z\ln(z+\xi)}{2(z+\xi)^7}+\frac{48z\ln 2}{(z+\xi)^7}+\frac{21z\ln z}{2(z+\xi)^7}\right]. \qquad (33)$$

Substituting the values of z and ξ, one gets

$$E_{eN}^{(2)}=-0.0540B. \qquad (34)$$

From Eqs. (25) and (32), remembering $I=J=\frac{3}{2}$, it then follows that

$$a_c=\frac{4\mu_N\mu_0E_{eN}^{(2)}}{3a_0^3IJ(2\pi\hbar)B}=-5.24\ \text{Mc/sec}, \qquad (35)$$

where $\mu_N=\gamma_NIh$ is the magnetic moment of the Li^7 nucleus and μ_0 is the Bohr magneton.

Since $u_{1s}(r)$ involves a single exponential for the Coulson-Duncanson function, one can obtain an exact analytic solution for the perturbation equation (16) for the EP method, with $\delta\psi_{iN}$ and $\mathcal{K}_N$ replaced by $\delta\psi_{iE}$ and $\mathcal{K}_E$, respectively.

$$\delta\psi_{1s,E}=\frac{8}{9}\left(\frac{z^3}{\pi}\right)^{1/2}\frac{\xi^5e^{-zr}}{(z+\xi)^7}\left\{\frac{3(z+\xi)^4r^2e^{-2\xi r}}{4\xi}+\left[\frac{3(z+\xi)^4}{4\xi^2}+\frac{9(z+\xi)^3}{2\xi}\right]re^{-2\xi r}+\left[\frac{3(z+\xi)^4}{8\xi^3}+\frac{9(z+\xi)^3}{4\xi^2}+\frac{51(z+\xi)^2}{4\xi}\right]e^{-2\xi r}+\frac{75e^{-2\xi r}}{4r}-\frac{24e^{(z-\xi)r}}{r}-\frac{75z}{2}\int\frac{e^{-2\xi r}}{r}dr+48z\int\frac{e^{(z-\xi)r}}{r}dr\right\}-\frac{28}{3}\left(\frac{z^3}{\pi}\right)^{1/2}\frac{z^2\xi^5e^{-zr}}{(z+\xi)^7}\left(r+\frac{1}{z}\ln r-\frac{1}{2z^2r}\right). \qquad (36)$$

Substituting $\delta\psi_{1s,E}$ from Eq. (36) into Eq. (12), one obtains, after some manipulation, exactly the same expression for $E_{eN}^{(2)}$ as in Eq. (33).

The value of a_c in Eq. (35) is about one-half of Goodings'[4] calculated value (−10.7 Mc/sec) by the SP method and is in somewhat better agreement with Cohen, Goodings, and Heine's[5] value (−8.5 Mc/sec) by the EP method. However, one does not expect very good agreement with these earlier results because the single-exponential hydrogen functions are not good approximations to the Hartree-Fock wave functions.[22]

To get a result by the MP method which could be properly compared in terms of accuracy with Cohen,

[20] C. Coulson and W. E. Duncanson, Proc. Roy. Soc. (Edinburgh) **62**, 37 (1944).

[21] C. Schwartz, Ann. Phys. (N. Y.) **6**, 156 (1959).

[22] Per-Olov Löwdin and K. Appel, Phys. Rev. **103**, 1746 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

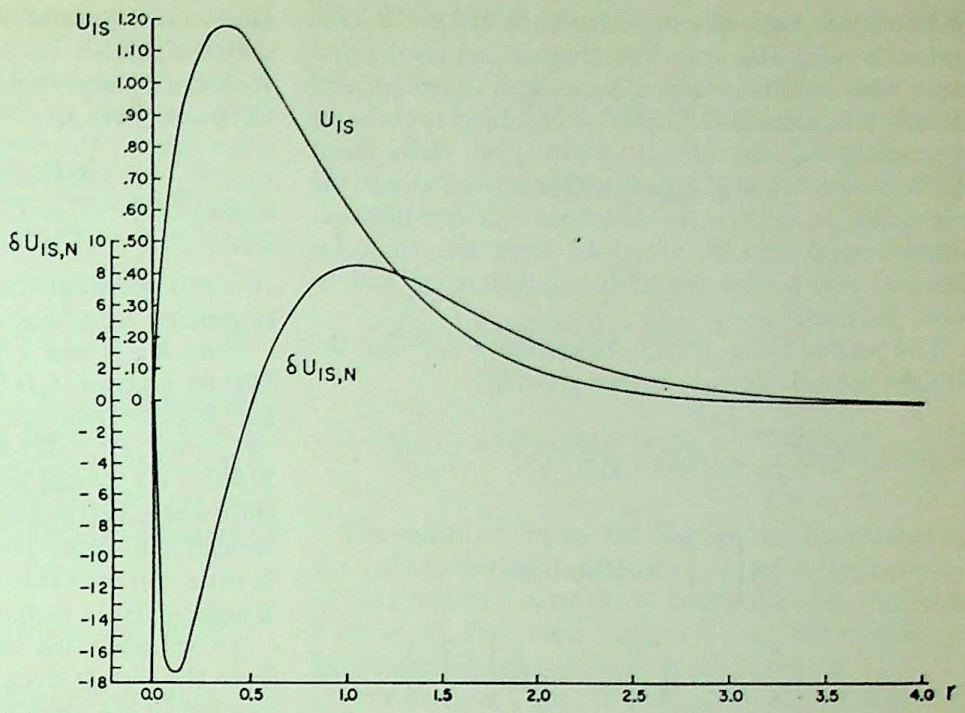


FIG. 1. Plot of unperturbed (u_{1s}) wave function and first-order perturbation $\delta u_{1s,N}$ as function of r for lithium atom $1s^22p$ configuration.

Goodings, and Heine's EP calculation and Goodings' SP calculation, we next solved Eq. (23) using Hartree-Fock wave functions. The solution has now to be obtained by numerical integration of Eq. (23). In Fig. 1, the function $\delta u_{1s,N}$ [defined in Eq. (32)] is plotted, together with Goodings' Hartree-Fock wave function[20] u_{1s}, which was employed in our calculations. The function $\delta u_{1s,N}$ resembles a $2s$ type function, in that, it has a node, but $\delta\psi_{1s,N}=(A/4\pi)(\delta u_{1s,N}/r)$ goes to negative infinity at the origin.

Using Eq. (31), together with Goodings' HF wave functions u_{1s}, u_{2p}, and $\delta u_{1s,N}$ of Fig. 1, it follows that

$$E_{eN}^{(2)}=-\tfrac{2}{3}B(0.137). \tag{37}$$

From Eq. (25) this leads to

$$a_c=-8.9 \text{ Mc/sec}. \tag{38}$$

This value compares very favorably with Cohen, Goodings, and Heine's EP value (Table I). This is satisfying in light of our Eqs. (12) and (13) and our earlier demonstration that EP and MP methods lead to identical results whenever exact solutions of Eqs. (9) and (10) are obtainable.

Both ours and Cohen, Goodings, and Heine's values for a_c are in slight disagreement with Goodings' value (Table I) obtained by the SP method. However, no measured value of a_c is available, so it is rather difficult to assess the importance of this small disagreement between the values obtained by MP (and EP) and SP methods.

The feasibility and accuracy of the MP method has now been demonstrated by its favorable comparison with the results of the EP method for the lithium atom $1s^22p$ state. We would next like to show the flexibility of the MP method by a consideration of core polarization effects for the $1s^22s$ state of lithium atom. The perturbation equation to be solved to get $\delta\psi_{1s,N}$ (or $\phi_{1s,N}$) in this case is exactly the same as in Eq. (26) except that u_{1s} may be different from its value for the $1s^22p$ state. Available tables[23] of u_{1s} for $1s^22s$ and $1s^22p$ states show that there is only a very slight difference between the values of u_{1s} in the two cases. One could

TABLE I. List of contributions to hyperfine constants in Mc/sec for lithium atom $1s^22s$ and $1s^22p$ states.

	$1s^22s$			$1s^22p$		
	Core polarization	Direct $2s$ contribution	Total	Core polarization	Direct $2p$ contribution	Total
Goodings	106	284	390	−10.7	0	−10.7
Cohen, Goodings, and Heine[a]	86.74	286.26	373.0	−8.5	0	−8.5
This paper	83.76	284	367.76	−8.9	0	−8.9
Experimental			401.786			

[a] The direct $2s$ contribution is taken from Table I of Ref. 5.

[23] D. A. Goodings, Ph.D. thesis, Cambridge University, Cambridge, England, 1961 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

therefore use the same $\phi_{1s,N}$ for both $1s^22p$ and $1s^22s$ states. If using this approximation we find good agreement with results obtained by Cohen, Goodings, and Heine's EP calculation (which is not subject to the same approximation), we can indeed justify our claim about the flexibility of the MP method. For heavier atoms, the alteration in core wave functions u_{ns} for different atomic configurations would be even less than for lithium; and so the flexibility approximation will be more justified.

The second-order energy expression $E_{eN}^{(2)}$ for the lithium atom $1s^22s$ state is then given by

$$E_{eN}^{(2)}=-2B\left\{\int_0^\infty u_{2s}(r)\delta u_{1s,N}(r)\times\left[\frac{1}{r}\int_0^r u_{1s}(s)u_{2s}(s)ds+\int_r^\infty u_{1s}(s)u_{2s}(s)\frac{1}{s}ds\right]dr\right\}. \quad (39)$$

On evaluating the integrals in Eq. (39), one gets $E_{eN}^{(2)}=2B(0.1434)$. The hyperfine constant is therefore

$$a_c=83.76 \text{ Mc/sec}. \quad (40)$$

The comparison of this value with Cohen, Goodings, and Heine's and Goodings' values (Table I) is very satisfying.

IV. CORE CONTRIBUTION TO KNIGHT SHIFT IN LITHIUM METAL

As another important application of the MP method, we shall consider the magnetic field produced at the Li^7 nucleus in lithium metal as a result of the exchange polarization of the core $1s$ electrons when the conduction electrons are polarized by the external magnetic field. The usual expression for the Knight shift due to the conduction electrons is given by[24]

$$\frac{\Delta H}{H}=\frac{8\pi}{3}\chi_p\langle|\psi_F(0)|^2\rangle_{\text{av}}, \quad (41)$$

where χ_p is the spin susceptibility per atom, and $\langle|\psi_F(0)|^2\rangle_{\text{av}}$ is the average electron density at the nucleus from electrons at the Fermi surface. Equation (41) can be interpreted in the following way: Of the conduction electrons near the Fermi surface, a fraction $(\frac{1}{2}+\chi_pH/2\mu_0)$ have their spins parallel to the field, and a fraction $(\frac{1}{2}-\chi_pH/2\mu_0)$ have their spins antiparallel to the field. When these electrons interact with the nuclear moment through the interaction term $\mathcal{H}_N$, we get the effective field ΔH in Eq. (41) out of the difference in populations in the parallel and antiparallel spin states. This population difference between the parallel and antiparallel conduction electron states also leads to an exchange polarization potential to the core electrons, given by

$$\mathcal{H}_E=-\frac{\chi_pH}{\mu_0}\frac{\psi_k(1)}{\psi_{\text{core}}(1)}\int\psi_k(2)\frac{1}{r_{12}}\psi_{\text{core}}(2)d\tau_2. \quad (42)$$

It is this potential $\mathcal{H}_E$ which enables the core electrons to produce an additional field at the nucleus.

Since $\mathcal{H}_E$ is now a weaker perturbation than in Eq. (14) by a factor χ_pH/μ_0, the perturbation methods EP and MP are more justified than in atoms. Also, it is clear from Eq. (42) that terms of the order $\mathcal{H}_E^2\mathcal{H}_N$, $\mathcal{H}_E^3\mathcal{H}_N$, $\cdots$, would lead to contributions to the Knight shift which are proportional, respectively, to the first, second, and higher powers of the field. However, there is little current evidence for such contributions to the Knight shifts in metals.

To calculate the contribution to the Knight shift from the core $1s$ electrons by the MP method, we then proceed in the same manner as in Sec. III to calculate $E_{eN}^{(2)}$ as given by Eq. (13). For the perturbed core wave function, the function $\delta u_{1s,N}$ (Fig. 1) for the atomic lithium $1s^22p$ state was used. This is again an approximation because there may be a small difference between the wave functions u_{1s} for the metal, and for the atomic $1s^22p$ state. However, our experience with the lithium atom $1s^22s$ state in Sec. III indicates that the error due to this approximation has little effect on $E_{eN}^{(2)}$. We would also like to point out that this error would be expected to be smaller for the metal than for the $1s^22s$ atomic state because the conduction electron wave functions appear to have predominantly p character at the Fermi surface.[17]

Making use of Eqs. (13), (25), and (42), and remembering that we have to average over the Fermi surface in calculating Knight shifts, we get

$$\frac{\Delta H}{H}=\frac{8\pi}{3}\chi_p\frac{2}{A_F}\left\langle\int_0^{r_s}\psi_{kF}(1)\delta\psi_{1s,N}(1)\times\psi_{kF}^*(2)\psi_{1s}(2)\frac{1}{r_{12}}d\tau_1d\tau_2\right\rangle_{\text{Fermi surface}}, \quad (43)$$

where ψ_{kF} is the wave function for a conduction electron at the Fermi surface, r_s is the radius of the Wigner-Seitz sphere, and A_F is the area of the Fermi surface. Equation (43) can be evaluated using available wave functions ψ_{kF} for lithium metal. Of the published wave functions for lithium, those which are available in a readily usable form are the recent ones by Kohn and Callaway.[17]

$$\psi_k=\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{(N)^{1/2}}[u_0(r)+iku_1(r)P_1(\cos\theta_{kr})+k^2(u_2(r)P_2(\cos\theta_{kr})+\phi(r))]. \quad (44)$$

[24] W. D. Knight, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1956), Vol. 2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The potential used in obtaining these results was an empirical potential which is expected to be more accurate than a calculated potential based on some model. In Eq. (44), θ_{kr} is the angle between a radius vector $\mathbf{r}$ and wave vector $\mathbf{k}$. The functions $u_0(r)$, $\phi(r)$, $u_1(r)$, and $u_2(r)$ were given in tabulated form and N is a normalization factor calculated up to k_F^2 term (where k_F is the Fermi momentum in the spherical Fermi surface approximation).

$$N=4\pi(1+2.79333k_F^2). \tag{45}$$

Since the Fermi surface in lithium approximates a sphere quite closely, we can take the magnitude of $\mathbf{k}_F$ in all directions as the same. The details of the calculations for evaluating the integral in Eq. (43) are given in the Appendix. It is shown there that including up to terms of order k_F^2 [Eq. (A4)]

$$\begin{aligned}
\psi_{k_F}(r)&=As(k_F,r)P_0(\cos\theta_{kr})+Ap(k_F,r)P_1(\cos\theta_{kr})\\
&\quad+Ad(k_F,r)P_2(\cos\theta_{kr}),\\
As(k_F,r)&=\frac{1}{(4\pi)^{1/2}}\Big[j_0(k_F,r)u_0(r)-k_Fj_1(k_F,r)u_1(r)\\
&\quad+k_F^2\Big(j_0(k_F,r)\phi(r)-j_2(k_F,r)u_2(r)\\
&\quad-\frac{2.79333}{2}j_0(k_F,r)u_0(r)\Big)\Big],\\
Ap(k_F,r)&=\frac{1}{(4\pi)^{1/2}}\Big[3u_0j_1+k_F(j_0u_1-2j_2u_1)\\
&\quad+k_F^2\Big(1.2j_1u_2+3j_1\phi-1.8j_3u_2\\
&\quad-\frac{3(2.79333)}{2}u_0j_1\Big)\Big],\\
Ad(k_F,r)&=\frac{1}{(4\pi)^{1/2}}\Big[-5j_2u_0+k_F(3j_3u_1-2j_1u_1)\\
&\quad+k_F^2\Big(j_0u_2-\frac{10}{7}j_2u_2-5j_2\phi+\frac{18}{7}j_4u_2\\
&\quad+\frac{5(2.79333)}{2}j_2u_0\Big)\Big],
\end{aligned}\tag{46}$$

where $j_n(k_F,r)$ is the spherical Bessel function of the order n. Using Eq. (46), it is shown in the Appendix [Eq. (A15)], that one obtains after some manipulation

$$\begin{aligned}
\frac{\Delta H}{H}&=\frac{8\pi\chi_p}{3}\frac{1}{2\pi}\int\delta\psi_{1s,N}(r_1)\psi_{1s}(r_2)\\
&\quad\times\Big\{(4\pi)^3\frac{As(k_F,r_1)As^*(k_F,r_2)}{r_>}\\
&\quad+3\Big(\frac{4\pi}{3}\Big)^3Ap(k_F,r_1)Ap^*(k_F,r_2)\frac{r_<}{r_>^2}\\
&\quad+5\Big(\frac{4\pi}{5}\Big)^3Ad(k_F,r_1)Ad^*(k_F,r_2)\frac{r_<^2}{r_>^3}\Big\}r_2^2dr_2r_1^2dr_1.
\end{aligned}\tag{47}$$

The contributions to $\Delta H/H$ from the three terms in Eq. (47) involving As, Ap, and Ad refer, respectively, to the s, p, and d parts of the conduction electron wave function at the Fermi surface. These were evaluated by numerical integration and are listed in Table II. It can be seen that $(\Delta H/H)_{cp}$ is almost equal to $(\Delta H/H)_{cs}$ but opposite in sign. In contrast, $(\Delta H/H)_{cd}$ is about a factor of 25 smaller than both $(\Delta H/H)_{cp}$ and $(\Delta H/H)_{cs}$. This leads us to expect that the contributions to the core polarization from f and higher angular momentum parts of the conduction electron wave function would be negligible. Unfortunately, it is not possible to test this point using Callaway and Kohn's wave function, because they did not give higher angular momentum components beyond d. The value of the total core contribution, $(\Delta H/H)_c$ in units of $(647.5\ \chi_p)$, namely -0.00544, is to be compared with the value of 0.10280 in the same units, for $(\Delta H/H)_{\text{direct}}$ that one gets using Kohn and Callaway's wave functions and Eq. (41). So we find that the core polarization contributes about -5.3% correction to the Knight shift. Also it is to be noted that, while the s character of the conduction electron wave function is alone important, for the direct contribution to the Knight shift as given by Eq. (41), the core polarization depends sensitively on the s and p parts of the wave function, since there is a near cancellation between $(\Delta H/H)_{cs}$ and $(\Delta H/H)_{cp}$. It is therefore imperative to know accurately the relative amounts of s and p characters in the wave functions for other metals, since the core-polarization contribution may be a more significant fraction of the direct contribution. This point has also been noted by Cohen, Goodings, and Heine.[5] These authors assumed that the conduction electron wave function can be well approximated by a combination of $2s$ and $2p$ atomic functions. In such a case, the core polarization can be calculated by taking a weighted combination of the core polarizations for atomic $1s^22s$ and $1s^22p$ configurations. However, in Figs. 2 and 3, we have compared As and Ap with the $2s$ and $2p$ atomic wave functions, respectively, obtained by Goodings.[4] It is seen that, while there is a reasonable

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

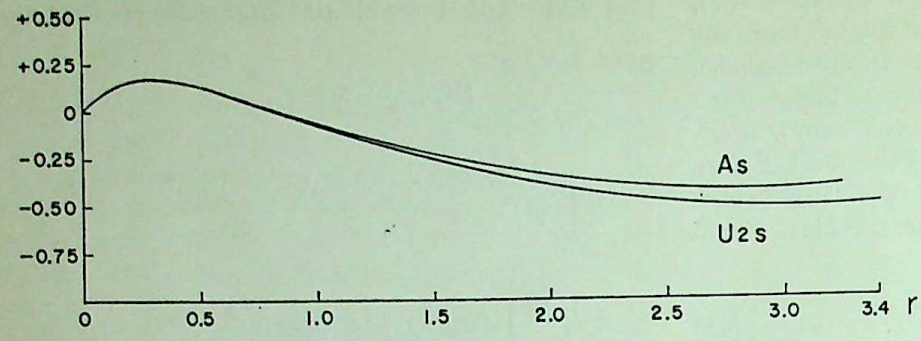


FIG. 2. Comparison of the s part of the conduction electron wave function (at Fermi surface) with Goodings' $2s$ wave function (u_{2s}) for lithium atom $1s^22p$ configuration.

similarity in shapes between the metallic and atomic wave functions, the correspondence is far from exact. To make a quantitative application of the EP method comparable to our MP calculations, one would have to solve two equations of the type of Eq. (9) using Eq. (42) for $\mathcal{H}_E$ and then calculate $(\Delta H/H)_c$ using Eq. (12). With the MP method, however, because of the flexibility in using $\delta\psi_{1s,N}$ as long as the core wave functions do not vary substantially, we did not have to solve any new differential equations to obtain $(\Delta H/H)_c$. This is an illustration of the computational advantage of the MP method.

In our numerical calculations leading to the results in Table II, we had to make use of the expressions in Eq. (46) correct to k_F^2. This was necessary because of the form of Eq. (44) in which Callaway and Kohn's wave function was available. However, Eq. (47) can be used for any general wave function (such as orthogonalized plane wave and augmented plane wave), since one can obtain As, Ap, and Ad by expanding the conduction electron wave function in spherical harmonics.

TABLE II. List of contributions to Knight shift in lithium metal.

Contributions[a] to $\Delta H/H$	Value
Core-polarization s part, $(\Delta H/H)_{cs}$	0.02355
Core-polarization p part, $(\Delta H/H)_{cp}$	−0.02813
Core-polarization d part, $(\Delta H/H)_{cd}$	−0.00085
Total core polarization $(\Delta H/H)_c$	−0.00544
Direct contribution[b] $(\Delta H/H)_{\text{direct}}$	0.10280
Grand total $(\Delta H/H)_{\text{total}}$	0.09736
Experiment[c]	0.101

[a] The tabulated numbers are in units of a_0^{-3} (atomic units). To correct to the usual dimensionless manner (%) that Knight shift data are obtained, multiply by 647.5 χ_p, where χ_p is the spin susceptibility per gram.
[b] Calculated using Kohn and Callaway's wave function (Ref. 17).
[c] Taken from Table IV, Ref. 24.

CONCLUSION

As we mentioned at the outset, our motivation in developing the MP method was to find a procedure for investigating the contribution from exchange polarization to Knight shifts in metals and alloys. While the EP method as developed by Cohen, Goodings, and and Heine can, in principle, be applied in any situation, one would like to have a method that is computationally simpler and flexible enough to apply for a nucleus in a variety of environments. Our experience with lithium nucleus in the three situations discussed here leads us to believe that the MP method can be extended to give reliable results in other cases. The necessity for careful investigations of core polarizations is strongly felt at the present time, because a variety of Knight shift data in metals, alloys, and internal fields in ferromagnets, are now available; and it appears that direct contributions from conduction electrons cannot quantitatively explain such data.

We would like to conclude by making a few comments on the self-consistency problem in core-polarization calculations in general and specifically as pertinent to the MP method. In an atom containing an unpaired electron, the many-electron wave function has to be

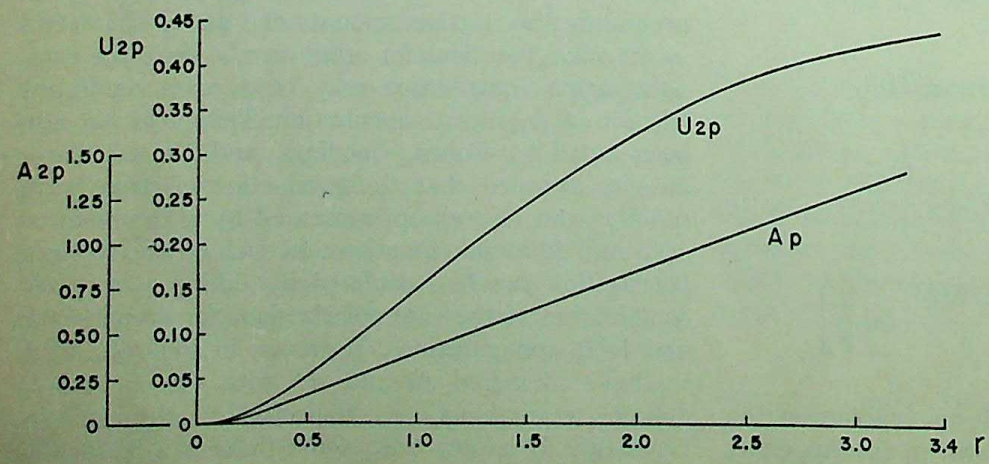


FIG. 3. Comparison of the p part of the conduction electron wave function (at Fermi surface) with Goodings' $2p$ wave function (u_{2p}) for lithium atom $1s^22p$ configuration.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

an eigenfunction of S^2. As has been pointed out by a number of authors, in a correct application of the Hartree-Fock method, this condition must be built in as a constraint. This is the case with the PSP method. The SP method, however, does not have this constraint, and the results from both SP or SPP methods are therefore somewhat suspect. Heine has pointed out that it is the self-consistency procedure in the SP method which is responsible for its dangers because it leads to an uncertain admixture of correlation and exchange effects and may in fact overemphasize the correlation unduly. The EP method, which considers the exchange polarization as a perturbation and has no self-consistency built in, is free from this danger although it is evidently less accurate than PSP. In the MP method, however, the perturbed state is not an eigenfunction of S^2 because of the nature of $\mathcal{H}_N$; hence, there is no condition of constraint regarding the total spin. Self-consistency can therefore be used in the MP method without any fear of spurious contributions from correlation. However, our aim was to extend this method to metals where self-consistency is rather difficult to apply even for restricted Hartree-Fock calculations. We have therefore not explored the effects of self-consistency (Dalgarnos' method I) here.

ACKNOWLEDGMENTS

The authors are grateful to Professor J. Callaway and Dr. G. Burns for helpful discussions. We would also like to acknowledge the help we received from Dr. M. J. Garber, Director of the Computing Services at the University of California, Riverside.

APPENDIX

From Eq. (24), the second-order change in energy due to the exchange polarization of the core electrons by the conduction electrons at the Fermi surface is

$$E_{eN}^{(2)}=-\frac{2}{4\pi k_F^2}\left\langle\int_0^{r_s}\psi_{k_F}(1)\delta\psi_{1s,N}(1)\times\psi_{k_F}^*(2)\psi_{1s}(2)\frac{1}{r_{12}}d\tau_1 d\tau_2\right\rangle_{\text{Fermi surface}},\quad\text{(A1)}$$

where r_s is the radius of the Wigner, Seitz sphere $=3.21\,a_0$, k_F is the Fermi momentum $=0.59788/a_0$, and

$$\psi_{k_F}=\frac{e^{i\mathbf{k}_F\cdot\mathbf{r}}}{(N)^{1/2}}[u_0(r)+ik_Fu_1(r)P_1(\theta_{kr})+k_F^2(u_2(r)P_2(\theta_{kr})+\phi(r))],\quad\text{(A2)}$$

θ_{kr} being the angle between $\mathbf{k}_F$ and $\mathbf{r}$, and N the normalization factor, $4\pi(1+2.79333k_F^2)$ and for abbreviation we denote $P_l(\cos\theta_{kr})$ by $P_{l(\theta_{kr})}$ where θ_{kr} is the angle between $\mathbf{k}_F$ and $\mathbf{r}$. On expanding $e^{i\mathbf{k}_F\cdot\mathbf{r}}$ and N, and keeping terms up to k_F^2, ψ_{k_F} can be written as

$$\psi_{k_F}(r)=As(k_F,r)P_0(\theta_{kr})+Ap(k_F,r)P_1(\theta_{kr})+Ad(k_F,r)P_2(\theta_{kr}),\quad\text{(A3)}$$

where

$$As(k_F,r)=\frac{1}{(4\pi)^{1/2}}\Big[j_0(k_F,r)u_0(r)-k_Fj_1(k_F,r)u_1(r)+k_F^2\Big(j_0(k_F,r)\phi(r)-j_2(k_F,r)u_2(r)-\frac{2.79333}{2}j_0(k_F,r)u_0(r)\Big)\Big],$$

$$Ap(k_F,r)=\frac{1}{(4\pi)^{1/2}}\Big[3j_1u_0+k_F(j_0u_1-2j_2u_1)+k_F^2\Big(1.2j_1u_2+3j_1\phi-1.8j_3u_2-\tfrac{3}{2}(2.79333)u_0j_1\Big)\Big],\quad\text{(A4)}$$

$$Ad(k_F,r)=\frac{1}{(4\pi)^{1/2}}\Big[-5j_2u_0+k_F(3j_3u_1-2j_1u_1)+k_F^2\Big(j_0u_2-\frac{10}{7}j_2u_2-5j_2\phi+\frac{18}{7}j_4u_2+\tfrac{5}{2}(2.79333)u_0j_2\Big)\Big],$$

where $j_n(k_F,r)$ is the spherical Bessel function of order n. Using Eqs. (A3) and (A4), Eq. (A1) takes the form

$$E_{eN}^{(2)}=-\frac{1}{2\pi}\int\delta\psi_{1s,N}(r_1)\psi_{1s}(r_2)\frac{1}{r_{12}}\{[As(k_F,r_1)+Ap(k_F,r_1)P_1(\theta_{k1})+Ad(k_F,r_1)P_2(\theta_{k1})]\times[As^*(k_F,r_2)+Ap^*(k_F,r_2)P_1(\theta_{k2})+Ad^*(k_F,r_2)P_2(\theta_{k2})]\}d\tau_1d\tau_2\sin^2\theta_{k_F}d\theta_{k_F}d\phi_{k_F},\quad\text{(A5)}$$

θ_{k1} representing the angle between $\mathbf{k}_F$ and $\mathbf{r}_1$ and θ_{k2} the angle between $\mathbf{k}_F$ and $\mathbf{r}_2$. We perform the integration over the spherical Fermi surface first, and then integrate out the angular parts of $d\tau_1$ and $d\tau_2$.

Since the directions of $\mathbf{r}_1$ and $\mathbf{r}_2$ are arbitrary, let $\mathbf{r}_1$ be directed along $\mathbf{k}_z$ for convenience. Using the addition theorem for spherical harmonics

$$P_2(\theta_{k2})=\frac{4\pi}{5}\sum_{m=-2}^{2}Y_2^{m*}(\theta_{k1},\phi_{k1})Y_2^m(\theta_{12},\phi_{12}),$$
$$P_1(\theta_{k2})=\frac{4\pi}{3}\sum_{m=-1}^{1}Y_1^{m*}(\theta_{k1},\phi_{k1})Y_1^m(\theta_{12},\phi_{12}),\quad\text{(A6)}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

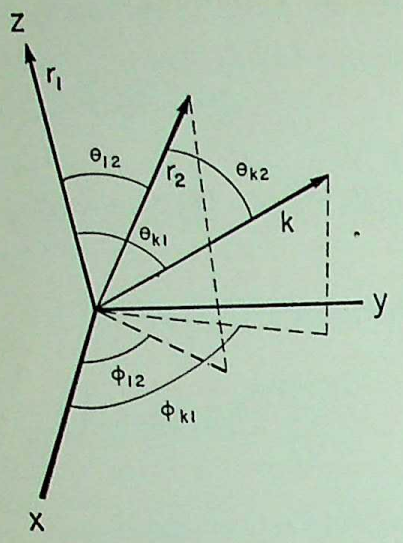


FIG. 4. Relative orientations of $\mathbf{k}$, $\mathbf{r}_1$, and $\mathbf{r}_2$.

one can evaluate the following integrals over the Fermi surface:

$$\frac{4\pi}{3}\int P_1(\theta_{k1})\sum_{m=-1}^{1} Y_1{}^{m*}(\theta_{k1},\phi_{k1})Y_1{}^{m}(\theta_{12},\phi_{12})d\Omega_k = \frac{4\pi}{3}P_1(\theta_{12})\delta_{m0}, \quad \text{(A7)}$$

$$\frac{4\pi}{5}\int P_2(\theta_{k1})\sum_{m=-2}^{2} Y_2{}^{m*}(\theta_{k1},\phi_{k1})Y_2{}^{m}(\theta_{12},\phi_{12})d\Omega_k = \frac{4\pi}{5}P_2(\theta_{12})\delta_{m0}, \quad \text{(A8)}$$

$$\int P_j(\theta_{k1})P_i(\theta_{k2})d\Omega_k = 0, \quad i\neq j.$$

Equation (A5) then reduces to

$$E_{eN}{}^{(2)} = -\frac{1}{2\pi}\int \delta\psi_{1s,N}(r_1)\psi_{1s}(r_2)\frac{1}{r_{12}}\Big[4\pi As(k_F,r_1)\times As^*(k_F,r_2)+\frac{4\pi}{3}Ap(k_F,r_1)Ap^*(k_F,r_2)P_1(\theta_{12}) +\frac{4\pi}{5}Ad(k_F,r_1)Ad^*(k_F,r_2)P_2(\theta_{12})\Big]d\tau_1 d\tau_2. \quad \text{(A9)}$$

In order to perform the angular integration involved in Eq. (A9) the following relationships will be used:

$$\frac{1}{r_{12}} = 4\pi\sum_{l=0}^{\infty}\frac{1}{(2l+1)}\frac{r_<{}^l}{r_>{}^{l+1}}\sum_{m=-l}^{l} Y_l{}^{m*}(\theta_1,\phi_1)\times Y_l{}^{m}(\theta_2,\phi_2), \quad \text{(A10)}$$

$$P_1(\theta_{12}) = \frac{4\pi}{3}\sum_{m=-1}^{1} Y_1{}^{m*}(\theta_1,\phi_1)Y_1{}^{m}(\theta_2,\phi_2), \quad \text{(A11)}$$

$$P_2(\theta_{12}) = \frac{4\pi}{5}\sum_{m=-2}^{2} Y_2{}^{m*}(\theta_1,\phi_1)Y_2{}^{m}(\theta_2,\phi_2),$$

where (A11) again follows from the spherical harmonic addition theorem. Thus,

$$\int\frac{1}{r_{12}}P_1(\theta_{12})d\Omega_1 d\Omega_2 = \frac{(4\pi)^2}{3}\sum_{l=0}^{\infty}\frac{1}{(2l+1)}\frac{r_<{}^l}{r_>{}^{l+1}} \times\sum_{m'=-l}^{l}\int Y_l{}^{m'*}(\theta_1,\phi_1)Y_l{}^{m'}(\theta_2,\phi_2) \times\sum_{m=-1}^{1} Y_1{}^{m}(\theta_1,\phi_1)Y_1{}^{m*}(\theta_2,\phi_2)d\Omega_1 d\Omega_2 = \left(\frac{4\pi}{3}\right)^2\frac{r_<}{r_>{}^2}\times 3. \quad \text{(A12)}$$

$$\int\frac{1}{r_{12}}P_2(\theta_{12})d\Omega_1 d\Omega_2 = \left(\frac{4\pi}{5}\right)^2\frac{r_<{}^2}{r_>{}^3}\times 5, \quad \text{(A13)}$$

$$\int\frac{1}{r_{12}}P_0(\theta_{12})d\Omega_1 d\Omega_2 = (4\pi)^2\frac{1}{r_>}. \quad \text{(A14)}$$

The final expression for $E_{eN}{}^{(2)}$ is then

$$E_{eN}{}^{(2)} = -\frac{1}{2\pi}\int \delta\psi_{1s,N}(r_1)\psi_{1s}(r_2) \times\{(4\pi)^2As(k_F,r_1)As^*(k_F,r_2)(1/r_>) +3(4\pi/3)^3Ap^*(k_F,r_2)Ap(k_F,r_1)(r_</r_>{}^2) +5(4\pi/5)^3Ad(k_F,r_1)Ad^*(k_F,r_2)(r_<{}^2/r_>{}^3)\} \times r_2{}^2dr_2 r_1{}^2dr_2, \quad \text{(A15)}$$

which involves only radial integrations over r_1 and r_2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Millimeter-Wave Molecular-Beam Spectroscopy: Alkali Chlorides*

PAUL L. CLOUSER† AND WALTER GORDY
Department of Physics, Duke University, Durham, North Carolina
(Received 20 December 1963)

The pure rotational spectra of the alkali chlorides were investigated in the 0.96- to 3-mm range of the microwave region with the molecular-beam spectrometer earlier developed at Duke University. Introduction of Teflon microwave lenses and high-pass microwave filters improved this spectrometer so that measurements into the submillimeter region were possible, to an accuracy of better than one part in 10^6. Dunham's solution for the diatomic molecule was applied in interpretation of data. Improved values for B_e, α_e, and γ_e were obtained for most molecules studied. The centrifugal distortion constants D_e and β_e were obtained from the rotational spectra for the first time for all molecules measured. From the latter two constants, accurate values of ω_e and $\omega_e x_e$ were derived. Other derived quantities are: potential coefficients, isotopic mass ratios, moments of inertia, and internuclear distances. For most of these quantities, the accuracies obtained surpass those from previous measurements.

INTRODUCTION

An earlier designed, high-temperature, molecular-beam, microwave spectrometer[1,2] for operation in the shorter millimeter-wave range of the spectrum has been improved and extended in frequency coverage and has been used for extensive measurements of the alkali chlorides. Unlike the rf molecular-beam spectrometers,[3–5] which detect deflected molecules, this spectrometer detects the loss in radiation which results from the absorption of radiation by a beam of molecules simply sprayed across the radiation path. Since no deflection of the beam is necessary, this spectrometer can be used at the very high millimeter- or submillimeter-wave frequencies, at which molecules generally become insensitive to the deflection fields required for operation of the molecular-beam electric-resonance spectrometers. It has the advantage over the latter of employing a microwave detector which does not depend on the type of molecules being investigated. Because of the weak absorption of molecules in the centimeter or longer wave rf region, the present instrument, which we shall designate as a molecular-beam absorption spectrometer, is not generally applicable at these frequencies. It becomes applicable and effective in the shorter millimeter- and submillimeter-wave regions, where the absorption coefficients of most molecules become large. Because molecules generally become less sensitive to electric and magnetic fields with increase of rotational frequency, the molecular-beam resonance methods become ineffective in the millimeter-wave region, where the present spectrometer operates best. Thus, the molecular-beam absorption spectrometer is complementary to the electric- and magnetic-resonance spectrometers.

Before development of our molecular-beam spectrometer, the microwave group at Columbia University designed a hot-cell microwave spectrometer[6,7] which, like ours and like the conventional microwave spectrometer for gases, employs a microwave radiation detector. With it, they measured rotational transitions in the centimeter wave region for most of the alkali halides. More recently, Lide, Cahill, and Gold[8] have improved the hot-cell spectrometer and have extended its range to 45 kMc/sec for the measurement of the $J=0\rightarrow 1$ transition of LiCl. The principal disadvantage of the hot-cell spectrometer is the pronounced pressure and Doppler broadening of the spectral lines in a gas heated to high temperature. Also, the microwave-transmission losses in a hot cell are large for the high millimeter-wave frequencies. The molecular-beam spectrometer used here eliminates most of the pressure and Doppler broadening, and its microwave-transmission losses are sufficiently low that it can be used in the submillimeter-wave region. Introduction of dielectric lenses for the focusing of the millimeter waves through the absorption cell has significantly reduced its transmission losses over those of the earlier design.

As was demonstrated in the study of alkali bromides and iodides[2] and as is evident from this study of the alkali chlorides, a precise measurement of the millimeter-wave rotational spectra allows a more complete and more accurate evaluation of the molecular constants of these molecules than is possible from either infrared or centimeter-wave measurements. The millimeter-wave measurements are needed because the effects of centrifugal distortion on the molecular structure can be significant and because these effects of distortion may be

* This study was supported by the U. S. Air Force Office of Scientific Research. Grant No. AF-AFOSR-493-64.
† Now at the Bell Telephone Laboratories, Holmdel, New, Jersey.
[1] A. K. Garrison and W. Gordy, Phys. Rev. 108, 899 (1957).
[2] J. R. Rusk and W. Gordy, Phys. Rev. 127, 817 (1962).
[3] H. K. Hughes, Phys. Rev. 72, 614 (1947).
[4] C. A. Lee, B. P. Fabricand, R. O. Carlson, and I. I. Rabi, Phys. Rev. 91, 1395 (1953).
[5] B. P. Fabricand, R. O. Carlson, C. A. Lee, and I. I. Rabi, Phys. Rev. 91, 1403 (1953).
[6] M. L. Stitch, A. Honig, and C. H. Townes, Rev. Sci. Instr. 25, 759 (1954).
[7] A. Honig, M. Mandel, M. L. Stitch, and C. H. Townes, Phys. Rev. 96, 629 (1954).
[8] D. R. Lide, P. Cahill, and L. P. Gold, J. Chem. Phys. 40, 156 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

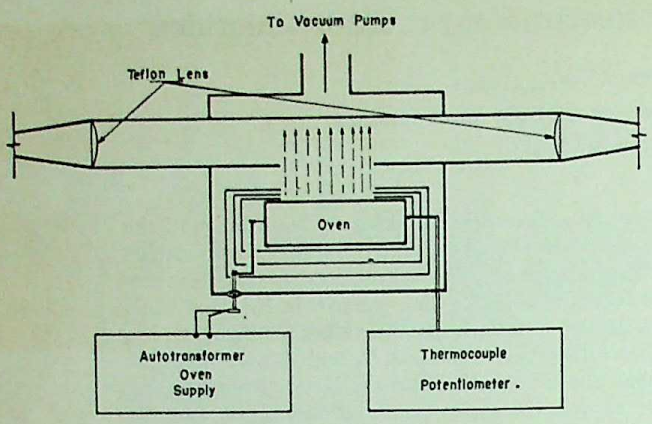


FIG. 1. Schematic diagram of molecular beam absorption cell.

accurately evaluated only by measurements at the higher microwave frequencies, at which frequencies this distortion significantly alters the rotational spectra.

EXPERIMENTAL CONSIDERATIONS

In its systems for microwave-power generation, detection, control, and measurement, the spectrometer used in this study is similar to most of those employed for video detection in the Duke microwave laboratory.[9] The principal difference between this spectrometer and conventional microwave spectrometers is in the absorption cell itself, which is specialized for the generation and utilization of high-temperature molecular beams.

Since the molecular-beam spectrometer has been described in earlier papers of this series,[1,2] only the present modifications of it will be described here. The absorption cell is shown in schematic arrangement in Fig. 1. It consists of a 60-cm section of *S*-band waveguide (3- $\times 1\frac{1}{2}$-in. cross section), into which the output of a harmonic generator operating from a 32–39 kMc/sec reflex klystron is matched by means of a 30-cm *G*- to *S*-band transition section.[9] The power is matched to the *G*-band waveguide of the detector by an identical *S*- to *G*-band transition section. A 6-in. slit is cut in the bottom of the *S*-band waveguide which is the absorption cell so that a beam of molecules can pass into it from a separate oven located just below the slit. The spectrometer has been modified from the earlier design[2] by the addition of high-pass microwave filters which reduce detector-crystal noise in order to give higher harmonics and the addition of Teflon lenses which increase the efficiency of transmission through the slotted waveguide section. These modifications have made possible detection into the submillimeter region. The filters, which operate by means of a short constriction in the *E*-field direction in an otherwise straight section of rectangular waveguide, are located at the output guide of the harmonic generator. Three high-pass microwave filters were made with cutoff wavelengths corresponding to the fourth, fifth, and sixth harmonics of the klystron frequency at the low end of its frequency band. These cutoff wavelengths were 1.87 mm, 1.56 mm, and 1.17 mm, respectively.

The lenses were placed at the large ends of the *G*- to *S*-band transition sections, as shown in Fig. 1. They are identical, planoconvex types designed upon assumption of a point source of radiation about one centimeter from the apex of the input transition section. Y. Y. Hu[10] shows that for sectoral *E*- or *H*-plane horns of less than a five-degree half-angle, the formula $d/\lambda = 5$ gives the location of the phase center or "equivalent point source" of the horn for the design of parabolic reflectors. That the theory should follow for transmission lenses is evident. The distance d into the horn from its apex for the wavelength for which the lenses were designed ($\lambda = 2$ mm) was approximately a centimeter and was not at all critical. The "equivalent point source" appears to be actually a sphere with a radius of about 1 cm for the small-angle horn used here, as is noted in Hu's work. However, with the lens 29 cm from the equivalent source, the point-source approximation is a good one.

By use of the equivalent point source, two lenses were designed by means of the thin-lens equation normally employed in the design of optical lenses. Since the wavelength of 2 mm is small as compared with the 76-$\times$38-mm lenses, the thin-lens equation is a reasonable approximation from consideration of diffraction effects. Teflon was used for fabrication of the lenses because of its high softening temperature. Its index of refraction was taken to be 1.35. A focal length of 29 cm causes the lens to collimate the output beam of the horn. This collimated beam of microwave power is refocused by its mate at the receiving horn so that it matches this power into the *G*-band waveguide of the detector. Brown's monograph on microwave lenses[11] offers further information on this subject.

The errors in measurement with the present spectrometer compare favorably with those obtained with the methods of molecular-beam resonance. The accuracy of measurement is limited only by the observer's ability to ascertain the center of an absorption line. Six measurements were made on each line, with directions of the electronic sweep of the klystron alternated so that time delays in the amplifier are canceled. This gives three measurements for which the dispersion usually lies between 1 and 5 parts in 10^7.

No very difficult problems were encountered in the measurement of the rotational transitions in the more abundant chlorides of Na, K, Rb, and Cs. Difficulties with line strength in most of the less abundant chlorine and rubidium isotopes precluded measurements over the three vibrational states necessary for resolution of B_v into the three structural constants Y_{01}, α_e, and γ_e.

[9] W. Gordy, W. V. Smith, and R. F. Trambarulo, *Microwave Spectroscopy* (John Wiley and Sons, New York, New York, 1953), Chap. I.

[10] Y. Y. Hu, J. Franklin Inst. 271, 31 (1961).

[11] J. Brown, *Microwave Lenses* (Methuen and Company Ltd., London, 1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The spectra of LiCl were not observed because of the high degree of dimerization in the molecular beam. Lide, Cahill, and Gold[8] have reported measurements of the $J=0\rightarrow 1$ transition of LiCl made with a heated Stark modulation cell.

Cesium chloride, the heaviest molecule investigated in this study has significant populations in the first four vibrational states. This division of population among vibrational states causes the rotational absorption lines for the ground vibrational state of this molecule to be weaker than those of the lighter chlorides, but allows measurement of rotational transitions in four of the vibrational states for $CsCl^{35}$.

All the isotopes of rubidium chloride studied have significant nuclear quadrupole couplings and at the lower J values show an absorption line splitting of several megacycles per second. Nevertheless, for the high J transitions measured in this work, the nuclear quadrupole splitting is negligible. For example, the line splitting is less than a tenth of a megacycle at 190 kMc/sec and is thus within the experimental error of our measurements. The rotational absorption lines in the third vibrational state were very weak; only one was measurable.

The spectral frequencies of $NaCl^{35}$ were measured over a wide range extending into the submillimeter-wavelength region. A cathode-ray oscilloscopic display of one of the submillimeter lines obtained for $NaCl^{35}$ is shown in Fig. 2. The signal is weak because it is near the upper frequency end of the operating range of the present spectrometer. Much stronger signals were, of course, observed in the one- to three-mm range.

Approximate oven temperatures at which the spectral lines appeared for each molecule are shown in Table I.

TABLE I. Oven temperatures.

Molecule	Temperature (°K)
NaCl	1022
KCl	1010
RbCl	970
CsCl	889

No attempt was made to measure these temperatures with strict accuracy. Care was taken only to obtain reproducible temperature measurements, since these measurements were used to predict the temperature at which spectral lines should appear at subsequent observations.

THEORETICAL CONSIDERATIONS

Dunham's theory[12] for the diatomic molecule has been applied in the analysis of the data obtained for the alkali halides. The part of this treatment which is required to make the derived results meaningful is outlined below.

[12] J. L. Dunham, Phys. Rev. 41, 713 (1932); 41, 721 (1932).

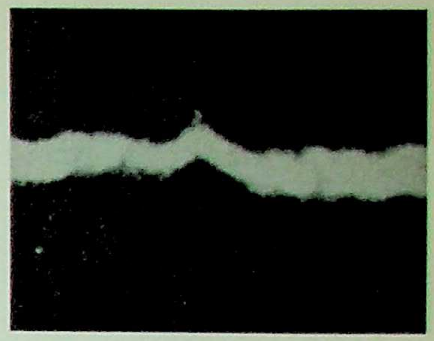


FIG. 2. Cathode-ray oscilloscopic display of the $J=23\rightarrow 24$, $v=0$, transition of $NaCl^{35}$ occurring at 312 109.88 Mc/sec (wavelength 0.97 mm). The width of the line is approximately 100 kc/sec.

The time-dependent, Schrödinger equation for the diatomic vibrating rotor may be written as

$$d^2\psi/d\xi^2+(2\mu r_e^2/h^2)\times[(E-U)-J(J+1)/(1+\xi)^2]\psi=0, \quad (1)$$

where $\xi=(r-r_e)/r_e$, where E is the energy value, and J is the rotational quantum number. The term $J(J+1)/(1+\xi)^2$ may be considered as a "centrifugal potential" and μ as a reduced mass in the equivalent, one-dimensional problem. The term U is dependent on the separation distance of the nuclei and represents the interatomic potential caused by the nuclear and electronic charges. For solution of this equation, the form of the potential function U must be known.

Many potential functions of the closed formula type have been examined by Varshni[13] and by other investigators.[14–16] Probably the most successfully and widely used of these functions is the Morse potential.[14] However, the best fit to the microwave data is obtained by use of a potential function expressed as a power series in ξ, valid only near a minimum and used first by Dunham.[12] This potential is

$$U(\xi)=hca_0^2\xi^2(1+a_1\xi+a_2\xi^2+a_3\xi^3+\cdots), \quad (2)$$

where the a's are constant coefficients.

Dunham applied this power series expansion of the potential in Eq. (1), expanded the centrifugal term as a power series, and solved the resulting equation about a minimum by the Wentzel-Kramers-Brillouin method.[12] The formula which he thus obtained for the energy levels is

$$F_{v,j}=\sum_{lj} Y_{lj}(v+1/2)^l J^j(J+1)^j, \quad (3)$$

where the Y_{lj} are the spectral constants to be obtained from the frequency measurements, J is the rotational quantum number, and v is the vibrational quantum number. If the units of the Y_{lj} are defined as Mc/sec, $F_{v,j}$ becomes a power series for the energy levels in the same units.

Dunham's approach has been reviewed by Sandeman[15] and has been extended slightly to include those Y_{lj}

[13] Y. P. Varshni, Rev. Mod. Phys. 29, 664 (1957).
[14] P. M. Morse, Phys. Rev. 34, 57 (1929).
[15] I. Sandeman, Proc. Roy. Soc. (Edinburgh) 60, 210 (1940).
[16] J. E. Kilpatrick, J. Chem. Phys. 30, 801 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

expressions which would be needed if measurements of the present accuracy could be made on vibrational transitions as well. Kilpatrick,[16] using a perturbation approach different from that of Dunham, has arrived at the same results as Sandeman and Dunham, in different form, except for his equation for Y_{00}, which includes D, a term for the dissociation energy. The term for the dissociation energy does not appear in the Dunham solution because this solution considers only the form of the potential in a restricted region near the minimum of the potential function. Sandeman has pointed out that the Dunham potential can be valid only over a limited region because of difficulty with convergence of the series potential for large ξ. Dunham's formulation, or Kilpatrick's, seems to be the only one, however, which can be used with the precise spectral data obtainable in the microwave region. It is satisfactory for the microwave measurements probably because the energy levels accessible to our microwave spectrometer lie sufficiently near the minimum of the potential curve for the formulation to apply. This theory was applied by Honig *et al.*[7] to their centimeter-wave measurements of the alkali halides made with their hot-cell spectrometer and by Rusk and Gordy[2] to their measurements of the alkali bromides and iodides made with their high-temperature molecular-beam spectrometer.

By applying the Bohr frequency condition,

$$\nu=(W_2-W_1)/h,$$

and the pure rotational absorption selection rules, $\Delta J=+1$, $\Delta v=0$, to the energy-level formula, one obtains the equation relating frequency ν, quantum numbers v, J, and structural constants Y_{ij}. The equation given below expresses this frequency in terms of those Dunham constants which have a measurable effect on the energy levels. The rotational quantum number for the lower level of the transition is represented by J.

$$\begin{aligned}\nu=&2Y_{01}(J+1)+2Y_{11}(v+\tfrac{1}{2})(J+1)\\&+2Y_{21}(v+\tfrac{1}{2})^2(J+1)+4Y_{02}(J+1)^3\\&+4Y_{12}(v+\tfrac{1}{2})(J+1)^3+Y_{03}(J+1)^3(J+2)^3.\end{aligned}\quad(4)$$

The relation of Y_{03} to other well-known structural constants is given by

$$Y_{03}=(2D_2/3\omega_e^2)(12B_e^2-\alpha_e\omega_e).\quad(5)$$

Subtraction of this term, in frequency units, from each measured line frequency ν, gives a corrected frequency ν'. The corrected frequencies ν_i' are used in the calculation of the five remaining constants with a least-square technique. The derived constants are calculated from the Y values with the following relations given by Dunham[12]:

$$Y_{10}=\omega_e=2(B_e^3/Y_{02})^{1/2},\quad(6)$$

$$a_0=\omega_e^2/4B_eC,\quad(7)$$

$$a_1=Y_{11}\omega_e/6B_e^2-1,\quad(8)$$

$$a_2=Y_{12}\omega_e^3/48B_e^4+2.375+2.25a_1+1.125a_1^2,\quad(9)$$

$$a_3=\omega_e^2Y_{21}/30B_e^3-1-2a_1+6a_2+2.6a_1a_2-1.5(a_1^2+a_1^3),\quad(10)$$

$$Y_{20}=-\omega_ex_e=-1.5B_e(a_2-1.25a_1^2),\quad(11)$$

$$Y_{00}=B_e(3a_2-1.75a_1^2)/8,\quad(12)$$

$$Y_{03}=16B_e^5(3+a_1)/\omega_e^4,\quad(13)$$

$$Y_{04}=-(64B_e^7/\omega_e^6)(13+9a_1-a_2+2.25a_1^2),\quad(14)$$

$$\beta_{01}=7.5+7a_1-4.5a_2+7.5a_3-11.5a_1a_2+5.25(a_1^2+a_1^3),\quad(15)$$

$$B_e=Y_{01}/(1+B_e^2\beta_{01}/\omega_e^2).\quad(16)$$

$$I_e=h/8\pi^2B_e,\quad(17)$$

$$\mu=M_1M_2/(M_1+M_2),\quad(18)$$

$$r_e=(I_e/\mu)^{1/2},\quad(19)$$

where B_e is set as equal to Y_{01} on the right-hand side of all equations except Eq. (16).

The probable error limits in the evaluated Y values were obtained by use of a statistical method[2] designed to produce a 99% confidence in the stated errors. The limits of error for the derived constants are based on these errors and have a 99% confidence level or better. The bulk of the data analysis was performed on the IBM-7072 digital computor available in the Duke University Digital Computing Laboratory.

Except for Y_{01}, the observed Dunham constants Y are related to the more familiar spectral constants for diatomic molecules, as follows:

$$\begin{aligned}&Y_{01}\simeq B_e \quad Y_{21}=\gamma_e \quad Y_{12}=-\beta_e\\&Y_{11}=-\alpha_e \quad Y_{02}=-D_e \quad Y_{03}=H_e.\end{aligned}\quad(20)$$

Similarly,

$$Y_{10}=\omega_e \quad Y_{20}=-\omega_ex_e.\quad(21)$$

The first of these relations, $Y_{01}\simeq B_e$, is only approximate, but the other relationships hold with sufficient accuracy for the results described here. For a further description of these spectral constants, see Herzberg's monograph.[17]

Because of the high J values involved in the transitions observed in the present work, no nuclear hyperfine structure was resolved. A negligible broadening of some of the lines by nuclear quadrupole interactions were the only hyperfine effects which could be observed. Thus we shall not outline the theory of nuclear interactions here. A treatment of these effects, which are resolvable at lower frequencies, is given by Honig *et al.*[7]

[17] G. Herzberg, *Molecular Structure and Molecular Spectra* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1959), 2nd ed., Vol. I, p. 144.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Observed and calculated microwave frequencies of the alkali halides.

Transition	Measured (Mc/sec)	Calculated (Mc/sec)
$NaCl^{35}$		
$v=0$, $J=7 \rightarrow 8$	104 189.74±0.10	104 189.66
$v=0$, $J=11 \rightarrow 12$	156 248.60±0.10	156 248.63
$v=0$, $J=13 \rightarrow 14$	182 262.89±0.10	182 262.86
$v=0$, $J=14 \rightarrow 15$	195 265.39±0.10	195 265.38
$v=0$, $J=16 \rightarrow 17$	221 260.15±0.10	221 260.11
$v=0$, $J=17 \rightarrow 18$	234 251.83±0.10	234 251.87
$v=0$, $J=20 \rightarrow 21$	273 202.04±0.15	273 202.04
$v=0$, $J=23 \rightarrow 24$	312 109.88±0.20	312 109.84
$v=1$, $J=13 \rightarrow 14$	180 907.78±0.10	180 907.81
$v=1$, $J=14 \rightarrow 15$	193 813.63±0.10	193 813.59
$v=1$, $J=17 \rightarrow 18$	232 509.95±0.10	232 509.89
$v=1$, $J=23 \rightarrow 24$	309 787.82±0.10	309 787.79
$v=2$, $J=14 \rightarrow 15$	192 371.05±0.10	192 371.02
$NaCl^{37}$		
$v=0$, $J=8 \rightarrow 9$	114 701.33±0.10	114 701.35
$v=0$, $J=14 \rightarrow 15$	191 091.62±0.10	191 091.57
$v=0$, $J=16 \rightarrow 17$	216 531.44±0.10	216 531.48
$v=1$, $J=8 \rightarrow 9$	113 858.03±0.10	113 857.97
$v=1$, $J=11 \rightarrow 12$	151 783.57±0.10	151 783.62
$v=1$, $J=14 \rightarrow 15$	189 686.13±0.10	189 686.11
KCl^{35}		
$v=0$, $J=12 \rightarrow 13$	99 929.54±0.10	99 929.52
$v=0$, $J=19 \rightarrow 20$	153 677.48±0.10	153 677.46
$v=0$, $J=24 \rightarrow 25$	192 023.49±0.10	192 023.48
$v=0$, $J=29 \rightarrow 30$	230 320.56±0.10	230 320.56
$v=0$, $J=34 \rightarrow 35$	268 558.98±0.15	268 558.94
$v=1$, $J=12 \rightarrow 13$	99 316.44±0.10	99 316.40
$v=1$, $J=24 \rightarrow 25$	190 844.49±0.10	190 844.53
$v=1$, $J=29 \rightarrow 30$	228 905.95±0.10	228 905.90
$v=2$, $J=24 \rightarrow 25$	189 670.52±0.10	189 670.51
KCl^{37}		
$v=0$, $J=21 \rightarrow 22$	164 220.00±0.10	164 220.01
$v=0$, $J=23 \rightarrow 24$	179 121.98±0.10	179 121.92
$v=0$, $J=35 \rightarrow 36$	268 363.84±0.15	268 363.80
$v=1$, $J=18 \rightarrow 19$	140 997.10±0.10	140 997.03
$v=1$, $J=23 \rightarrow 24$	178 038.18±0.10	178 038.10
$Rb^{85}Cl^{35}$		
$v=0$, $J=33 \rightarrow 34$	177 967.57±0.10	177 967.52
$v=0$, $J=34 \rightarrow 35$	183 187.48±0.10	183 187.53
$v=0$, $J=41 \rightarrow 42$	219 690.65±0.10	219 690.65
$Rb^{85}Cl^{35}$		
$v=0$, $J=48 \rightarrow 49$	256 120.42±0.10	256 120.44
$v=1$, $J=20 \rightarrow 21$	109 440.74±0.10	109 440.77
$v=1$, $J=33 \rightarrow 34$	177 045.49±0.10	177 045.50
$v=1$, $J=34 \rightarrow 35$	182 238.38±0.10	182 238.39
$v=1$, $J=41 \rightarrow 42$	218 551.58±0.10	218 551.62
$v=1$, $J=48 \rightarrow 49$	254 791.36±0.10	254 791.47
$v=2$, $J=33 \rightarrow 34$	176 126.30±0.10	176 126.30
$v=2$, $J=34 \rightarrow 35$	181 292.02±0.10	181 292.14
$v=3$, $J=33 \rightarrow 34$	175 209.82±0.10	175 209.90
$v=3$, $J=34 \rightarrow 35$	180 348.70±0.10	180 348.78
$Rb^{87}Cl^{35}$		
$v=0$, $J=29 \rightarrow 30$	156 025.06±0.10	156 025.05
$v=0$, $J=33 \rightarrow 34$	176 777.46±0.10	176 777.45
$v=0$, $J=40 \rightarrow 41$	213 046.83±0.10	213 046.82
$v=1$, $J=27 \rightarrow 28$	144 890.88±0.10	144 890.87
$v=1$, $J=37 \rightarrow 38$	196 490.65±0.10	196 490.64
$Rb^{85}Cl^{37}$		
$v=0$, $J=34 \rightarrow 35$	176 196.16±0.10	176 196.15
$v=0$, $J=35 \rightarrow 36$	181 216.30±0.10	181 216.30
$v=0$, $J=41 \rightarrow 42$	211 311.17±0.10	211 311.15
$v=1$, $J=34 \rightarrow 35$	175 300.90±0.10	175 300.90
$v=1$, $J=35 \rightarrow 36$	180 295.47±0.10	180 295.470
$v=2$, $J=34 \rightarrow 35$	174 408.33±0.10	174 408.260
$CsCl^{35}$		
$v=0$, $J=35 \rightarrow 36$	155 062.73±0.10	155 062.72
$v=0$, $J=44 \rightarrow 45$	193 699.77±0.10	193 699.74
$v=0$, $J=53 \rightarrow 54$	232 250.96±0.10	232 250.92
$v=1$, $J=44 \rightarrow 45$	192 790.44±0.10	192 790.44
$v=1$, $J=53 \rightarrow 54$	231 159.53±0.10	231 159.46
$v=1$, $J=62 \rightarrow 63$	269 425.24±0.15	269 425.25
$v=2$, $J=44 \rightarrow 45$	191 883.04±0.10	191 883.03
$v=2$, $J=53 \rightarrow 54$	230 070.34±0.10	230 070.29
$v=3$, $J=44 \rightarrow 45$	190 977.56±0.10	190 977.53
$CsCl^{37}$		
$v=0$, $J=42 \rightarrow 43$	177 224.90±0.10	177 224.89
$v=0$, $J=59 \rightarrow 60$	246 912.71±0.10	246 912.69
$v=1$, $J=42 \rightarrow 43$	176 411.12±0.10	176 411.12
$v=1$, $J=46 \rightarrow 47$	192 760.59±0.10	192 760.52
$v=1$, $J=59 \rightarrow 60$	245 776.61±0.10	245 776.61

ROTATIONAL CONSTANTS

The measured frequencies for each of the alkali chlorides studied are listed in Table II. In the same table are given for comparison the theoretical values of the frequencies as calculated with the spectral constants given in Table III.

The six constants (Y_{01}, α_e, γ_e, D_e, β_e, and H_e) were calculated directly from the measurements for the isotopic species $CsCl^{35}$, $RbCl^{35}$, KCl^{35}, $NaCl^{35}$, and $Rb^{85}Cl^{37}$. For the species $CsCl^{37}$, KCl^{37}, $NaCl^{37}$, and $Rb^{87}Cl^{37}$, these measurements yielded only four constants (B_0, B_1, D_e, β_e) because observations of rotational transitions for only two vibrational states were made. For the latter, Y_{01} and α_e were derived from B_0 and B_1 by use of a value of γ_e calculated with the isotopic mass ratio formulas. These six structural constants for each molecule are shown in Table III.

Comparison of the RbCl and KCl data with previous results[4] obtained with molecular-beam electric resonance shows that the three constants making up $B_v(Y_{01},\alpha_e,\gamma_e)$ have comparably small errors for the two types of measurement and that they agree well within these errors. This experiment shows no improvement in B_v over that obtained from low J transitions accurately measured with molecular-beam, electric-resonance techniques because the fractional error in the measured frequencies for both experiments is comparable.

A comparison of the present data for CsCl and NaCl with those of the previously made hot-cell measurements[7] shows a significant improvement in accuracy in the values for the three constants making up B_v. The greatest improvement of the present study over the previous results arises in the measurement of the centrifugal distortion constants D_e and β_e and in the evaluation of other constants which are influenced by these constants. The earlier microwave measurements,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Observed rotational constants.

Molecule	Y_{01} Mc/sec	α_e $-Y_{11}$ Mc/sec	γ_e Y_{21} Mc/sec	D_e $-Y_{02}$ Kc/sec	β_e $-Y_{12}$ Kc/sec	H_e Y_{03} cps$\times 10^{-3}$
$NaCl^{35}$	6537.367±0.006	48.711±0.009	0.154±0.008	9.3540±0.0050	−0.0250±0.0050	−1.020 ±0.010
$NaCl^{37}$	6397.287±0.006	47.151±0.009	0.147±0.008	8.9630±0.0060	−0.0190±0.0060	−0.950 ±0.010
KCl^{35}	3856.373±0.009	23.681±0.013	0.049±0.004	3.2600±0.0030	−0.0025±0.0003	−0.623 ±0.008
KCl^{37}	3746.578±0.009	22.678±0.013	0.047±0.004	3.0780±0.0040	−0.0041±0.0040	−0.571 ±0.008
$Rb^{85}Cl^{35}$	2627.393±0.004	13.599±0.004	0.021±0.001	1.4830±0.0010	0.0007±0.0004	−0.248 ±0.002
$Rb^{87}Cl^{35}$	2609.774±0.004	13.458±0.006	0.020±0.001	1.4610±0.0010	0.0019±0.0010	−0.243 ±0.002
$Rb^{85}Cl^{37}$	2526.856±0.004	12.826±0.005	0.019±0.002	1.3710±0.0010	0.0003±0.0010	−0.221 ±0.002
$CsCl^{35}$	2161.246±0.002	10.119±0.002	0.011±0.002	0.9791±0.0004	0.0015±0.0002	−0.1413±0.0006
$CsCl^{37}$	2068.813±0.002	9.476±0.003	0.010±0.002	0.8974±0.0005	0.0014±0.0005	−0.1239±0.0007

which were performed at lower frequencies, did not yield these constants.

The slight effects of Y_{03} or H_e on the observed frequencies were obtained from values of the other structural constants with the equation

$$Y_{03}=H_e=2D_e(12B_e^2-\alpha_e\omega_e)/3\omega_e^2). \quad (22)$$

Only one pass through the computor was required for reduction of the measured frequencies to structural constants when this calculated value was used. The values of H_e as calculated from previous constants and as calculated from the constants derived from the first passage through the computor agreed within the limits of the effect of the H_e's on the measured frequencies. Values for H_e, listed in Table III, show about two more significant figures than would have been obtained if the value for H_e had been extraced directly from the measured frequencies.

The measured constants under consideration have been found to be consistent under the mass ratio transformation within the experimental errors. The values for Y_{00} and Y_{04} (Table IV) have a number of significant figures and thus provide some check for the Dunham theory. Derived values for Y_{00} and Y_{04} are consistent for different isotopic species of the same molecule, as are all the derived constants presented in the following sections.

Equation (16) was used for derivation of B_e from Y_{01}. The equilibrium moment of inertia and equilibrium nuclear distance, I_e and r_e, were derived from B_e by use of Eqs. (17) and (19), respectively. These values are listed with two-figure errors to allow for adequate comparison between isotopes. If, at a later date, a more accurate value of Planck's constant h should become available, these values may be calculated from the data with greater absolute accuracy. Values for I_e and r_e both agree within their experimental errors with previous measurements because the five-figure accuracy of h masks the differences between these and the hot-cell measurements. All conclusions reached by Honig *et al.*[7] regarding I_e and r_e remain unaffected. Values for these constants are listed in Table V.

TABLE IV. Derived values of Y_{00} and Y_{04}.

Molecule	Y_{00} (kMc/sec)	Y_{04}
$NaCl^{35}$	2.30±0.20	−0.003 ±0.002
$NaCl^{37}$	2.10±0.20	−0.004 ±0.002
KCl^{35}	1.29±0.08	−0.0010 ±0.0002
KCl^{37}	1.40±0.40	−0.0006 −0.0007
$Rb^{85}Cl^{35}$	0.73±0.08	−0.00033−0.00005
$Rb^{87}Cl^{35}$	0.80±0.20	−0.00040±0.00010
$Rb^{85}Cl^{37}$	0.80±0.20	−0.00026±0.00009
$CsCl^{35}$	0.40±0.04	−0.00020±0.00002
$CsCl^{37}$	0.40±0.10	−0.00017±0.00003

TABLE V. Molecular structural constants.

Molecule	B_e (Mc/sec)	β_{01}	I_e[a] (amu-Å)	r_e[a] (Å)
$NaCl^{35}$	6537.406±0.006	−17±2	77.3376±0.0030	2.360898±0.000046
$NaCl^{37}$	6397.320±0.006	−15±3	79.0311±0.0031	2.360898±0.000046
KCl^{35}	3856.385±0.009	−15±1	131.1040±0.0053	2.666772±0.000054
KCl^{37}	3746.592±0.009	−18±7	134.9459±0.0054	2.666771±0.000054
$Rb^{85}Cl^{35}$	2627.397±0.004	−10±2	192.4290±0.0076	2.786865±0.000055
$Rb^{87}Cl^{35}$	2609.776±0.004	−6±4	193.7282±0.0077	2.786869±0.000055
$Rb^{85}Cl^{37}$	2526.860±0.004	−12±5	200.0852±0.0079	2.786864±0.000055
$CsCl^{35}$	2161.247±0.002	−4±1	233.9332±0.0091	2.906411±0.000057
$CsCl^{37}$	2068.814±0.003	−3±4	244.3852±0.0096	2.906408±0.000057

[a] $h=(1.05443\pm0.00004)\times10^{-27}$ erg-sec, 1 g $=(6.02486\pm0.00016)\times10^{23}$ amu, from J. W. Dumond and E. R. Cohen, in *Handbook of Physics*, edited by E. U. Condon and H. Odishaw (McGraw-Hill Book Company, Inc., New York, 1958), pp. 7–168.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

VIBRATIONAL CONSTANTS

The vibrational constants ω_e and $\omega_e x_e$ were derived from the experimentally observed Y's by use of the Dunham theory, Eqs. (6)–(20). Values thus obtained are compared in Table VI with the values obtained from infrared measurements by Rice and Klemperer.[18] The present values of ω_e agree within experimental error with those from their infrared data, but are an order of magnitude more accurate if one can trust the Dunham theory to hold within the limits of our experimental error. It is worthy of note that the ω_e values of Rice and Klemperer are in much better agreement with our ω_e values, and also with those of Rusk and Gordy on the alkali bromides and iodides than are any other previous values from infrared or optical spectroscopy. Through an oversight their excellent infrared values were not listed for comparison in the earlier paper.[2]

TABLE VI. Derived vibrational constants of alkali chlorides.

Molecule	ω_e Microwave[a] cm^{-1}	ω_e Infrared[b] cm^{-1}	$\omega_e x_e$ Microwave[a] cm^{-1}	$\omega_e x_e$ Infrared[b] cm^{-1}
$NaCl^{35}$	364.60±0.10	366±4	1.755±0.030	2.05
$NaCl^{37}$	360.60±0.10		1.740±0.030	
KCl^{35}	279.80±0.10	281±6	1.167±0.005	1.30
KCl^{37}	275.80±0.20		1.113±0.005	
$Rb^{85}Cl^{35}$	233.34±0.08	228±6	0.856±0.008	0.92
$Rb^{87}Cl^{35}$	232.67±0.08		0.870±0.020	
$Rb^{85}Cl^{37}$	228.88±0.08		0.810±0.040	
$CsCl^{35}$	214.22±0.05	209±6	0.740±0.004	0.75
$CsCl^{37}$	209.55±0.06		0.710±0.020	

[a] From present data calculated with Dunham's theory.
[b] Infrared results of Rice and Klemperer (Ref. 18).

There is an interesting regularity evident in the comparison of the theoretical values of $\omega_e x_e$ from the Dunham equations, from a formula derived by Pekeris,[19] and the experimental values of Rice and Klemperer.[18] Pekeris' formula, shown below, gives better agreement with the $\omega_e x_e$ values of Rice and Klemperer than does the Dunham expression for the lighter molecules, particularly NaCl (see Table VII). The agreement of infrared data with the Dunham or Pekeris formula shows a regular variation with increasing molecular mass.

TABLE VII. Comparison of microwave and infrared $\omega_e x_e$'s for alkali chlorides and bromides.

Molecule	$\omega_e x_e$ Microwave[a] cm^{-1}	$\omega_e x_e$ Infrared[b] cm^{-1}	$\omega_e x_e$ Microwave[c] cm^{-1}
$NaCl^{35}$	1.755	2.05	2.0640
$NaBr^{79}$	1.160[d]	1.50	1.4040[d]
KCl^{35}	1.167	1.30	1.3390
NaI	0.965[d]	1.08	1.0720[d]
$Rb^{85}Cl^{35}$	0.856	0.92	0.9530
KBr^{79}	0.758[d]	0.80	0.8534[d]
$CsCl^{35}$	0.740	0.75	0.7940

[a] Calculated with Dunham's Theory.
[b] Infrared measurements, Rice and Klemperer (Ref. 18).
[c] Calculated with Pekeris's theory (Morse potential).
[d] Rusk and Gordy (Ref. 2).

[18] S. A. Rice and W. Klemperer, J. Chem. Phys. 27, 573 (1957).
[19] C. L. Pekeris, Phys. Rev. 45, 98 (1934).

The Pekeris formula for $\omega_e x_e$ follows from one for α_e, which was derived from a perturbation treatment with the Morse potential[14] as a basis.

$$\alpha_e = (2\omega_e x_e B_e/\omega_e)[3(B_e/\omega_e x_e)^{1/2} - 3(B_e/\omega_e x_e)]. \quad (23)$$

This equation may be solved for $\omega_e x_e$ to give

$$\omega_e x_e = B_e[(\alpha_e \omega_e/6B_e^2)+1]^2, \quad (24)$$

which in terms of Dunham's constant a_1 is

$$\omega_e x_e = B_e a_1^2. \quad (25)$$

Formula (25) may also be obtained from Dunham's relations if Y_{00} is equated to zero in Eq. (12) and the expression thus obtained for a_2 is substituted into Dunham's Eq. (11) for Y_{20}. By Eq. (11), Y_{20} equals $-\omega_e x_e$. The constant Y_{00} is invariant between isotopic species and is independent of the nuclear mass of a given molecule. Therefore, Y_{00} must be dependent on the nature of the potential function. It cannot arbitrarily be set to zero because its value is determined by the potential coefficients a_1 and a_2, whose values depend on the measured rotational constants. Table IV shows that the value for Y_{00} obtained from the rotational constants is definitely not zero.

The only spectral quantity which shows any possibly significant discrepancy within the Dunham theory is the anharmonicity constant, $\omega_e x_e$. This possible discrepancy should be examined further. Although Rice and Klemperer did not give estimates of the limits of error for their $\omega_e x_e$ values, the good agreement between their ω_e values and those of the present study gives credence to their $\omega_e x_e$ values. The Dunham theory is based upon an expansion of the potential function in a power series around the equilibrium position. The Morse potential function upon which the Pekeris formula is based should give a better representation than does the Dunham potential for regions not so near the equilibrium position. This might explain the better agreement of the Pekeris formula with the $\omega_e x_e$ values from the infrared measurement of the lighter molecules with their higher vibrational energies. Varshni and Sukla[20] compare these and other observed constants of the alkali halides with values calculated from a number of different potential functions. It is of interest to compare their different values for $\omega_e x_e$ (see their Table IV) with those obtained here.

MASS RATIOS

The mass ratios between Cl^{35} and Cl^{37}, also those between Rb^{85} and Rb^{87}, were calculated for all the molecules measured. All mass ratios are shown in Table VIII, along with some previously obtained values for

[20] Y. P. Varshni and R. C. Shukla, J. Chem. Phys. 35, 582 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VIII. Isotopic mass ratios.

Molecule	Chlorine (Cl^{35}/Cl^{37}) Method	Ratio	Molecule	Rubidium (Rb^{87}/Rb^{85}) Method	Ratio
NaCl	Present	0.9459780±0.0000045	$RbCl^{35}$	Present	0.9770086±0.0000112
KCl	Present	0.9459783±0.0000089	RbI	Present[d]	0.9770162±0.0000093
RbCl	Present	0.9459768±0.0000047	RbCl	MBER[e]	0.9770163±0.0000045
CsCl	Present	0.9459785±0.0000028	RbF	MBER[e]	0.9770148±0.0000052
CsCl	Hot Cell[a]	0.9459781±0.0000030	RbI	Hot cell[a]	0.9770177±0.0000045
$K^{39}Cl$	MBER[b]	0.9459803±0.0000015	RbBr	Hot cell[a]	0.9770146±0.0000055
	Mass spectroscopy[c]	0.9459759±0.0000035	RbBr	Mass spectroscopy[c]	0.9770191±0.0000022

[a] See Ref. 7.
[b] L. R. Maxwell, S. B. Hendricks, and V. M. Mosley, Phys. Rev. 52, 968 (1937).
[c] W. H. Johnson, K. S. Quisenberry, and A. O. Nier, *Handbook of Physics*, edited by E. U. Condon and H. Odishaw (McGraw-Hill Book Company, Inc., New York, 1958), pp. 9–55.
[d] See Ref. 2.
[e] See Ref. 4.

comparison. These quantities were calculated from the ratios B_e/B_e', as described by Honig *et al.*[7]

POTENTIAL COEFFICIENTS

The coefficients in the Dunham potential series, Eq. (2) up to and including a_3, are listed in Table IX

TABLE IX. Potential coefficients.

Molecule	a_0 cm^{-1}	a_1	a_2	a_3
$NaCl^{35}$	152 400±90	−3.076±0.001	6.47±0.09	−11.0±1.0
$NaCl^{37}$	152 320±100	−3.076±0.001	6.39±0.10	−11.0±1.0
KCl^{35}	152 100±100	−3.226±0.002	6.96±0.04	−12.0±0.8
KCl^{37}	152 100±200	−3.226±0.003	7.10±0.30	−13.0±3.0
$Rb^{85}Cl^{35}$	155 300±100	−3.297±0.001	7.08±0.08	−11.5±0.8
$Rb^{87}Cl^{35}$	155 500±100	−3.297±0.002	6.90±0.20	−10.0±2.0
$Rb^{85}Cl^{37}$	155 400±100	−3.297±0.002	7.14±0.20	−12.0±2.0
$CsCl^{35}$	159 100±70	−3.317±0.002	6.92±0.04	−10.0±0.8
$CsCl^{37}$	159 090±90	−3.318±0.001	6.89±0.20	−10.0±2.0

for each individual molecule. They were obtained with Eqs. (6)–(20). The potential coefficients between isotopic species are equal (within the stated errors) as should be expected.

Table X, which includes results for the bromides and iodides by Rusk and Gordy,[2] compares values of a_0, a_1, a_2, and a_3 for three different halides of the same alkali metals. The coefficients a_1 are shown to be approximately the same between the iodide, the bromide, and the chloride for a given alkali metal, except for cesium, where the relationship is one of slightly increasing absolute value of a_1 with decreasing mass. The coefficients a_0 also show a regularity in that a_0 increases smoothly with decreasing halide mass for the same alkali metal. These regularities were noted in the results on the iodides and bromides previously reported from this laboratory.[2] This chloride work extends and confirms them. They had not been shown before the development of our millimeter-wave, molecular-beam spectrometer, apparently because values of D_e and β_e were not available to sufficient accuracy for precise calculation of the potential coefficients. For comparison, we have given in parentheses the potential coefficients earlier obtained by Honig *et al.*[7] It will be noted that in many instances there are rather wide discrepancies between our values and theirs.

TABLE X. Comparison of potential coefficients.[a]

Molecule	a_0 $cm^{-1}\times10^{-5}$	a_1	a_2	a_3
LiBr	1.42 (1.02)	−2.71 (−2.45)	6.8 (5.5)	−19 (−14)
LiI	1.35 (1.14)	−2.70 (−2.56)	5.0 (6.0)	−8 (−15)
NaCl	1.52 (1.66)	−3.07 (−3.14)	6.4 (9.2)	−11 (−31)
NaBr	1.48 (1.64)	−3.05 (−3.16)	6.5 (7.4)	−12 (−16)
NaI	1.42 (1.74)	−3.02 (−3.23)	5.9 (8.8)	−8 (−26)
KCl	1.52 (1.81)	−3.22 (−3.43)	7.0 (10.0)	−12 (−32)
KBr	1.48 (1.63)	−3.24 (−3.35)	6.9 (9.1)	−11 (−26)
KI	1.43 (1.64)	−3.25 (−3.41)	6.9 (9.1)	−11 (−25)
RbCl	1.55 (2.08)	−3.30 (−3.66)	7.0 (11.8)	−11 (−42)
RbBr	1.53 (1.73)	−3.33 (−3.49)	7.3 (10.3)	−13 (−33)
RbI	1.44 (1.64)	−3.34 (−3.49)	7.2 (10.2)	−11 (−32)
CsCl	1.59 (2.00)	−3.32 (−3.59)	7.0 (10.6)	−10 (−35)
CsBr	1.55 (2.03)	−3.38 (−3.72)	7.6 (11.7)	−14 (−41)
CsI	1.50 (1.53)	−3.43 (−3.45)	7.6 (9.2)	−13 (−26)

[a] Data for bromides and iodides come from Rusk and Gordy (Ref. 2), those for chlorides from the present study. Values given in parentheses are those of Honig *et al.* (Ref. 7).

FUTURE WORK

Modifications of the oven are being made which will permit measurements to be made on substances which vaporize at higher temperatures than do the alkali chlorides. Millimeter-wave rotational transitions of NaF in different vibrational states have already been detected,[21] and it is hoped that a series similar to these on bromides, iodides, and chlorides can soon be reported from this laboratory on the alkali fluorides.

[21] S. E. Veazey and W. Gordy (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Precision Evaluation of the High-Frequency Limit of the Continuous X-Ray Spectrum Using a Gas Target X-Ray Tube*†

JON J. SPIJKERMAN‡ AND J. A. BEARDEN
Department of Physics, The Johns Hopkins University, Baltimore, Maryland
(Received 30 December 1963)

The major uncertainty in previous measurements of the high-frequency limit of the continuous x-ray spectrum was due to the fine structure associated with the observed cutoff. The development of a high-power mercury vapor target x-ray tube permitted observations to be made under idealized thin-target conditions. The x-ray intensity near the high-frequency limit was comparable to that of a solid target x-ray tube. The x-ray intensity was measured as a function of the voltage applied across the tube, with a double-crystal monochromator adjusted for the 1.537400 kxu wavelength of the Cu $K\alpha_1$ line. The absence of fine structure in the observed isochromat, in contrast to most solid targets, greatly increased the accuracy in the determination of the voltage cutoff. Accurate measurements of this voltage yielded a voltage-wavelength conversion factor $V\lambda_s = (12372.26 \pm 23$ ppm) xu-volts. Using h/e from non-x-ray data, the x-ray wavelength conversion factor from x units to cm (or angstroms) can be calculated, or conversely h/e evaluated. The effects of a revision in the Cu $K\alpha_1$ wavelength, which is now in publication by one of the authors, are also discussed.

INTRODUCTION

ELEMENTARY quantum theory considerations give the short wavelength limit of the continuous x-ray spectrum in terms of the voltage applied to the x-ray tube. This relation is $Ve = h\nu$, which may be rewritten as

$$h/e = V\lambda_s(\Lambda/c^2), \tag{1}$$

where h/e (the ratio of Planck's constant to the charge on the electron) is expressed in erg·sec/esu, V in volts, λ_s in kxu, and c, the velocity of light, in cm/sec. $\Lambda = \lambda_g/\lambda_s$, the x-ray wavelength conversion factor, represents the ratio between wavelength on the Siegbahn scale and wavelength on the absolute scale, i.e., Å/kxu. (1 kxu $= 10^3$ xu.)

This relationship has traditionally been used as a means of measuring h/e, but of course the x-ray measurement only evaluates $V\lambda_s$. The three[1–3] most precise evaluations of $V\lambda_s$ and the calculated values of h/e have yielded values which were considerably lower than those obtained through non-x-ray measurements.[4] The discrepancy was so great that DuMond and Cohen[5] gave such values a low weight in their evaluation of atomic constants while Bearden and Thomsen[4] excluded these measurements entirely. This discrepancy was largely attributed to the fine structure observed at the high-frequency limit of the continuous spectrum. Detailed measurements of the structure as a function of atomic number and voltage applied to the x-ray tube have been made by Johnson[6] and Per Johansson.[7] Its origin has been interpreted by Ulmer[8] as due to solid-state band structure. The initial state of the electron is determined by the cathode of the x-ray tube while the final state is characterized by the target. The distribution of those unoccupied energy levels to which a significant number of transitions are allowed then determines the shape of the continuous spectrum near cutoff. In principle this result holds only for a "thin target"; however since characteristic energy losses[9] occur in discrete steps of the order of 10 eV, it will also hold for a thick target within a few electron volts of the cutoff point.

In order to eliminate the fine structure, due to the solid state, a gas target x-ray tube[10] has been developed in which an electron beam of 20–50 mA and 8–20 kV is focused to a mercury jet as target. The intensity observed near the high-frequency limit is comparable to that of a solid target. This effectively thin target satisfies the assumptions of the Sommerfeld[11] theory and hence greatly simplifies the interpretation of the measurements which can then be used to obtain a refined value of h/e. On the other hand, if h/e is known with sufficiently high accuracy from non-x-ray sources, it gives a means of determining Λ, as is evident from Eq. (1). At the present moment, the relative uncertainty in the value of Λ appears somewhat greater than

* Research supported by U. S. Atomic Energy Commission Contract No. AT(30-1)-2543.
† Based in part on a thesis (J. J. S.) submitted to the Department of Physics, The Johns Hopkins University, 1962.
‡ Present address: National Bureau of Standards, Section 2.4, Washington, D. C. 20234.
[1] J. A. Bearden and G. Schwarz, Phys. Rev. 79, 674 (1950).
[2] J. A. Bearden, F. T. Johnson, and H. M. Watts, Phys. Rev. 81, 70 (1951).
[3] G. N. Felt, J. N. Harris, and J. W. M. DuMond, Phys. Rev. 92, 1160 (1953).
[4] J. A. Bearden and J. S. Thomsen, Suppl. Nuovo Cimento 5, 267–360 (1957); Am. J. Phys. 27, 569 (1959).
[5] J. W. M. DuMond and E. R. Cohen, Rev. Mod. Phys. 25, 691 (1953); E. R. Cohen, J. W. M. DuMond, T. W. Layton, and J. S. Rollett, *ibid.* 27, 363 (1955).
[6] F. T. Johnson, dissertation, Johns Hopkins University, Baltimore, Maryland, 1952 (unpublished).
[7] P. Johansson, Arkiv Fysik 18, 329 (1960).
[8] K. Ulmer, Z. Physik 159, 443 (1960); 162, 254 (1961); Phys. Rev. Letters 3, 514 (1959).
[9] L. Marton, L. B. Leder, and H. Mendlowitz, Advan. Electron. Electron Phys. 7, 183 (1955).
[10] W. Duane, Proc. Natl. Acad. Sci. (U. S.) 13, 662 (1927); 14, 450 (1928).
[11] A. Sommerfeld, Ann. Physik 11, 257 (1931).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

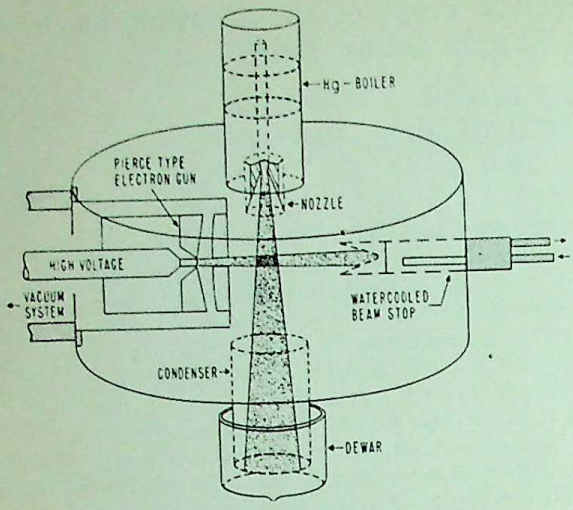


FIG. 1. Three-dimensional view of the gas target x-ray tube.

that of h/e, although both are of the same order of magnitude.

In interpreting the results, certain small corrections to the measured potential are essential, including those due to space charge and to work functions of the cathode and anode.[12] These are obtained, in part by a measurement of the ionization potential of the gas.

GAS TARGET X-RAY TUBE

Two gas target x-ray tubes have been used in the present work. The basic principles of both tubes are shown schematically in Fig. 1. The x rays were observed along an axis which was perpendicular to both the electron beam and the mercury jet.

The required area density of the jet can be estimated as follows: From solid target measurements[13,14] and electron energy losses in thin foils[9] it can be shown that the radiation near the high-frequency limit is produced by a solid target thickness of the order of 500 Å. Thus, to obtain sufficient x-ray intensity for precision measurements, the total atomic cross section for the gas target should be equivalent to a 500 Å thickness of a solid. This was achieved by using a vapor jet formed by a nozzle, as shown in Fig. 1. The diameter of the mercury beam as indicated by the focal spot was approximately 5 mm. The vapor was formed by an electrically heated Hg boiler, and the jetstream directed into a dry-ice cold trap. The freezing out of the Hg vapor greatly reduced the pumping speed required to maintain a pressure of 10^{-4} mm of Hg in the target chamber. A 500 liter/sec mercury vapor diffusion pump and liquid nitrogen traps were used to evacuate the electron gun and high voltage section of the x-ray tube.

The electron beam was formed by a two-element Pierce electron gun,[15] with a unipotential dispenser cathode.[16] This gun operated very satisfactorily, and at 8 kV produced a maximum beam current of 50 mA with a beam diameter of about 3 mm. The accelerating anode of the Pierce gun, the electron collector (beam stop), the nozzle, and the condenser were all electrically connected and maintained at ground potential, except for possible small effects such as contact and thermal emf's. The presence of Hg atoms in the tube largely neutralized the space charge[17] and collimated the electron beam through gas focusing. However, the ion bombardment of the cathode reduced its life to approximately 40 h.

In order to obtain effective work function corrections, as described in the section on corrections and probable errors, the ionization potential of Hg was measured by the usual method. This required a shielding grid at the potential of the target chamber and an ion collecting plate at a small negative potential. (These auxiliary electrodes are not shown in Fig. 1.) In the first tube a small grid and plate were inserted through the port where the x-ray beam was normally observed. In the second tube, both the stainless steel grid and plate were cylindrical and concentric with the mercury jet, with openings to permit passage of the electron and x-ray beams; they remained in place throughout the x-ray measurements, permitting an ionization curve to be run after each isochromat.

The first tube[18] was essentially two intersecting stainless steel cylinders each 2 in. in internal diameter and 4.5 in. long. The mercury jet was directly vertically upwards and an arrangement provided for condensing the mercury and returning it to the boiler. The second tube was designed on a larger scale, the actual geometry being very similar to the schematic shown in Fig. 1. The enclosing steel cylinder was 11 in. i.d. and 3.5 in. in height, with end plates of 0.5 in. thickness. The mercury jet was directly vertically downwards; no return was provided, since less than 300 cc of mercury was required for a 12-h run.

POWER SUPPLY AND MONOCHROMATOR

Power supply. The high-voltage regulated power supply and the high-voltage measurement circuit were basically the same as reported by Bearden and Schwarz,[1] with minor modifications. The regulating amplifier was replaced by a commercial high-gain operational amplifier, and a phase-shift-free voltage divider[19] was designed for the feedback loop. This circuit reduced the ripple voltage to 0.05 V peak to peak.

The high voltage V_m (see Table I) was measured with

[12] J. W. M. DuMond and V. L. Bollman, Phys. Rev. 51, 412 (1937).
[13] D. L. Webster, Phys. Rev. 9, 220 (1917).
[14] W. W. Nicholas, J. Res. Natl. Bur. Std. 2, 837 (1929).
[15] J. R. Pierce, *Theory and Design of Electron Beams* (D. Van Nostrand Company, Inc., New York, 1954), p. 152 ff.
[16] R. Levi, J. Appl. Phys. 26, 639 (1955).
[17] B. H. Wadia, J. Electron. Control 6, 307 (1959).
[18] J. J. Spijkerman, dissertation, Johns Hopkins University, Baltimore, Maryland, 1962, p. 12 (unpublished).
[19] C. L. Conner, Electronics 35, 52 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Measurements for the voltage-wavelength conversion factor.

	Run 1	2	3	4
		High-voltage measurement		
E (St. cell)	1.018 414	1.018 414	1.018 414	1.018 414
ΔE ($K-3$ pot.)	0.018 256	0.018 191	0.018 231	0.018 278
$E-\Delta E$	1.000 158	1.000 223	1.000 183	1.000 136
V_m	8048.96	8049.48	8049.16	8048.78
		Corrections		
V_i	10.434	10.434	10.434	10.434
V_{mi}	10.32	10.57	10.57	10.05
V_s-V_{si}	0.90	0.92	0.83	0.83
V_D	0.07	0.07	0.07	0.07
V	8048.10	8048.35	8048.12	8048.26
		Wavelength		
λ_s(xu)	1537.400	1537.370	1537.395	1537.403
		Voltage-wavelength conversion factor		
$V\lambda_s$[a]	12 373.14	12 373.29	12 373.13	12 373.41
Weight	1	2	2	2

[a] Weighted average $V\lambda_s = (12\,373.26 \pm 23$ ppm) xu·V.

a precision voltage divider and a potentiometer. The precision voltage divider, consisting of an 8526-Ω protective resistor in series with two 1-MΩ standard resistors (each with 10 subdivisions) and two standard resistors of 100 Ω, was so arranged that the smaller voltage was almost exactly equal to that of a standard cell for the average value of high voltage used in the measurements. The precision potentiometer then was used only to measure the small incremental difference between the voltage drop in the two 100-Ω standards of the divider and that of the standard cell. Hence the high voltage V_m was determined from the difference between the emf of the NBS 968 saturated standard cell and the potentiometer reading ΔE, multiplied by the voltage divider ratio of 8047.688. The standard resistors and standard cells used have been calibrated by the National Bureau of Standards a number of times, including one check during the present experiment.

Monochromator. A double-crystal spectrometer,[20] with quartz crystals cut parallel to the $10\bar{1}1$ plane, was used for a monochromator with a resolution of approximately 13 000 as determined from the parallel $(1, -1)$ position rocking curve. The monochromator was adjusted for the wavelength of the Cu $K\alpha_1$ line, which is one of the best known standard x-ray wavelengths. The monochromator was aligned by replacing the vapor target with a Cu target, and accurately calibrated against the $K\alpha_1$ line, assuming a peak wavelength of 1.537400 kxu.[21] In the different experimental runs the exact wavelength setting was determined with respect to this peak angle position. The x rays were detected by a scintillation counter with a 25-mil-thick NaI(Tl) crystal placed on a selected photomultiplier tube and shielded by a low radioactive iron housing. For a pulse-height discriminator setting which gave 80% detection efficiency the background was 0.3 counts/min.

ISOCHROMATS

In principle the high-frequency limit of the continuous spectrum can be found most simply by holding a fixed voltage on the x-ray tube and determining the shortest wavelength emitted. In practice it is usually more convenient to adjust the monochromator for a fixed wavelength and determine the variation of intensity as the voltage is changed in small steps. The resulting curve is called an isochromat.

Four averaged isochromats were used in the four runs shown in Table I. These were obtained as follows: Each averaged isochromat is the composite of ten individual isochromats. Seventeen voltage points in the vicinity of cutoff were selected and the x-ray intensity recorded for a period of 200 sec at each point. In order to minimize instrumental drift, these measurements were first recorded for increasing voltage and then repeated for decreasing voltage for each individual isochromat. After initial adjustment of the apparatus was completed and definitive recording was begun, no individual isochromat was rejected for any reason. One of the averaged isochromats is indicated by the data points in Fig. 2, where the normalized intensity is plotted against voltage change. Measurements up to 500 V above the short wavelength limit indicated no further appreciable increase in intensity.

Since the profile near the high-frequency limit is of primary interest, both for an h/e determination (or $V\lambda_s$) and from a theoretical standpoint, the observed

[20] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, Inc., New York, 1954), 2nd ed., pp. 709–740.

[21] J. A. Bearden and C. H. Shaw, Phys. Rev. 48, 18 (1935).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

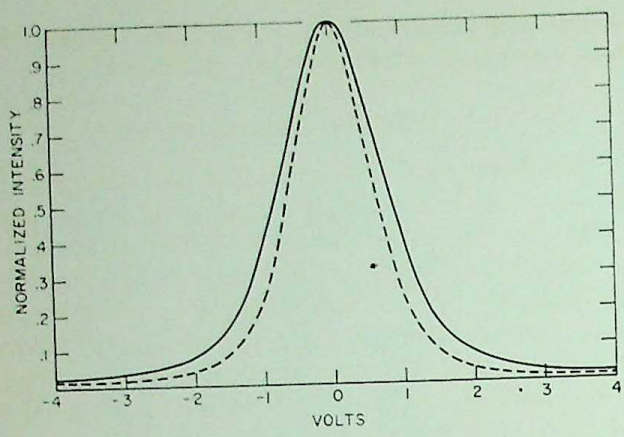


FIG. 2. Experimental window (solid line) and theoretical window (dashed line) of the double-crystal monochromator.

isochromat was corrected for instrumental smearing, which is primarily due to the width of the monochromator window. This window, shown in Fig. 3, was obtained from the parallel position (1,−1) rocking curve using a copper target. The double crystal diffraction pattern was also calculated from the Darwin-Prins-Ewald theory of single crystal diffraction.[22] The greater width of the observed pattern must be attributed to slight imperfections in the quartz crystals. A study of the crystals[23] showed that they were not perfect specimens. Since a 3 cm width of the crystal was used for both copper and vapor targets (due to the extended x-ray source), the crystal irregularities broadened the diffraction pattern.

The measured window was used to correct the observed isochromat, but the extended wings made an unfolding procedure inaccurate. This difficulty was avoided by using an IBM-7094 computer to fold assumed isochromat profiles with the observed window and finding the best fit for the observed data. The assumed profiles were based on the Sommerfeld theory,[24] which gives the x-ray intensity as a constant independent of the frequency for $h\nu<eV$, and zero for $h\nu>eV$. However, near cutoff the shape of the profile is determined by the distribution of final states available to the electron. If one assumes no electric or magnetic field at the target, the density of final states is proportional to $(\Delta E)^{1/2}$, where $\Delta E=e(V-V_0)$ or the energy difference between the incoming electron and the emitted photon.[25]

Consequently the x-ray intensity was represented by the empirical relation

$$I=I_0\{1-\exp[-\alpha(V-V_0)^{1/2}]\}, \qquad (2)$$

where α is an adjustable parameter. Equation (2) approximates the Sommerfeld profile, except near the high-frequency limit, and becomes proportional to the density of the final states as this limit is approached. The observed x-ray intensity should be given by folding Eq. (2) with the observed window of Fig. 3.

The parameters V_0 and α were varied stepwise; for each pair of values an isochromat was computed, and compared with the observed curve. The best fits were obtained with values of α between 1 and 2 V^{-1}; outside this range the predicted isochromat varied significantly from the observed one. The isochromats for these two values of α are shown in Fig. 2. For the best fit the value of V_0 differed only about 5 ppm from the voltage at the half-intensity point; hence, the latter was used as a sufficiently accurate criterion for the high-frequency cutoff.

CORRECTIONS AND PROBABLE ERRORS

Work Function Correction

The voltage measured, V_m, represents the potential across the x-ray tube and not the electron energy, since contact potentials and the initial thermal electron energy must be taken into account. In the solid target h/e experiments, the applied potential must be corrected for the work function of the cathode only; the work function of the anode is not involved, since the electron comes to rest near the Fermi level of the anode. For the gas target tube, however, both initial and final states are those of a free electron, and the difference between the work function of the cathode φ_c and the anode φ_a plus the thermal energy kT must be added to the measured potential.

The retardation of the electrons due to the space charge *at the cathode* has the effect of increasing the cathode work function. This space charge potential V_s

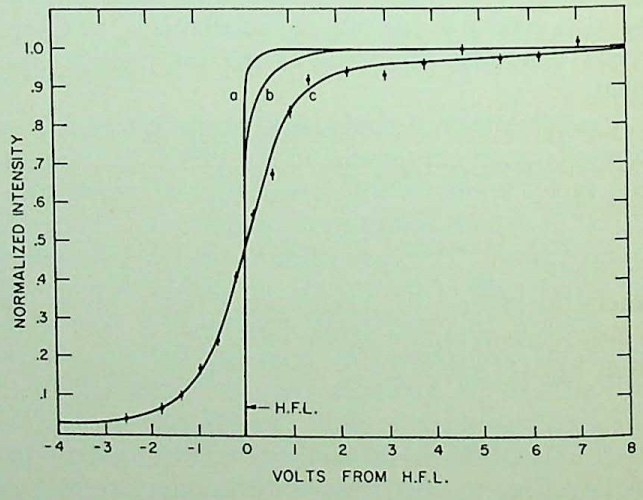


FIG. 3. (a) Assumed isochromat, Eq. (2), with $\alpha=1$ V^{-1}; (b) assumed isochromat with $\alpha=2$ V^{-1}; (c) predicted isochromat obtained by folding either assumed curve with the experimental window of Fig. 2; the predicted curves based on (a) and (b) showed no significant difference. Experimental data points are shown as dots, with bars indicating the statistical errors.

[22] The authors are indebted to J. G. Marzolf, S. J. for performing these calculations. See J. G. Marzolf, S. J., Bull. Am. Phys. Soc. 8, 313 (1963).
[23] A. Henins and J. A. Bearden, Bull. Am. Phys. Soc. 8, 313 (1963).
[24] A. Sommerfeld, Proc. Natl. Acad. Sci. (U. S.) 15, 393 (1929).
[25] E. P. Wigner, Phys. Rev. 73, 1002 (1948).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

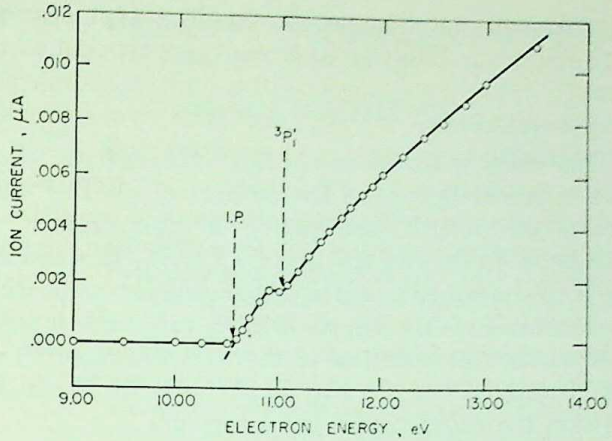


FIG. 4. Experimentally determined probability of ionization by electron impact in Hg gas. Cathode temperature 1493°K.

can be calculated[26] from the cathode work function φ_c, the critical cathode temperature θ (where the space charge minimum disappears), and operating temperature T, by

$$V_s=\varphi_c\left(\frac{T-\theta}{\theta}\right)-\frac{kT}{e}\ln\frac{I}{I_\theta}. \quad (3)$$

In this expression I_θ represents the cathode current at the critical temperature θ, and I that at operating temperature T, while k and e are the Boltzmann constant and the electronic charge. With the above-mentioned corrections to the measured voltage V_m, the effective difference V becomes

$$V=V_m+V_{kT}+(\varphi_c-\varphi_a)+V_s. \quad (4)$$

To determine these correction terms experimentally, a measurement was made of the ionization potential[27] of Hg (for which a value of 10.434 V is known from spectroscopic data). The observed probability of ionization by electron impact for Hg vapor is indicated in Fig. 4. The structure in the curve is due to autoionization [28] corresponding to the $^3P_1{}^1$ state of mercury 0.6 eV above the ground state, which competes with the ionization process.

For the ionization potential of mercury at the same cathode temperature, the relation corresponding to Eq. (4) is

$$V_i=V_{mi}+V_{kT}+(\varphi_c-\varphi_a)+V_{si}, \quad (5)$$

where the subscript i denotes quantities involved in the ionization potential measurement. Subtracting Eq. (5) from Eq. (4) and making use of Eq. (3) now yields

$$V=V_m+V_i-V_{mi}-\frac{kT}{e}\ln\frac{I}{I_i}, \quad (6)$$

the last term being simply V_s-V_{si}. Numerical values of these correction terms are shown in Table I.

Potential Depression

Since there is no external electric field acting on the electron beam in the target chamber, the space charge effects could produce a potential variation across the electron beam, which might become significant at moderate current densities.[29] The space charge effect would lower the electron energy. It should also produce an energy spread of several electron volts in the cutoff of the isochromat, unless it is largely neutralized by the presence of positive ions, which are trapped by the electron beam. Linder and Hernqvist[30] have studied the degree of neutralization of space charge by positive ions, and derived an expression for the potential depression as a function of the gas pressure.

For an operating pressure of 10^{-5} mm Hg in the described apparatus, the calculated voltage depression V_D is 0.07 V (Table I). This could be in error by 0.07 V, since complete neutralization would take place at 2×10^{-4} mm Hg and the pressure measurement could be in error by an order of magnitude.

Errors

The probable errors, in parts per million (ppm) of the final result are estimated as follows:

1. Determination of the voltage at the isochromat cutoff:	
Potentiometer measurements	3.5 ppm
Precision voltage divider	2.0
Isochromat cutoff	12
Work function correction	13
Voltage depression	9
Total (rms)	21 ppm
2. Spectrometer calibration	10 ppm

Total probable error in the voltage-wavelength conversion factor is thus 23 ppm; the error in the conversion factor from NBS to absolute volts (5 ppm) does not significantly increase this figure.

The four separate runs were assigned weights based on the statistics of the isochromat recording.

RESULTS

A summary of the experimental results is given in Table I. The total recording time for each run was slightly less than 19 h. Runs 1 and 2 were made with the tube briefly described in the gas target x-ray tube section and reported in detail in the thesis of one of the authors (J. J. S.). Runs 3 and 4 were made with the tube design shown in Fig. 1. The excellent agreement of the final values lends support to the methods and experimental measurements involved in applying the necessary corrections.

[26] Wayne B. Nottingham, in *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1957), Vol. 21, p. 42.
[27] S. C. Brown, *Basic Data of Plasma Physics* (John Wiley & Sons, Inc., New York, 1959), p. 14.
[28] W. M. Hickam, Phys. Rev. **95**, 704 (1954).
[29] O. Klemperer, *Electron Optics* (Cambridge University Press, London, 1953), p. 210.
[30] E. G. Linder and K. G. Hernqvist, J. Appl. Phys. **21**, 1088 (1950).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The voltage of standard cell E (actually three were used and frequently intercompared) has been corrected for the difference between the National Bureau of Standards volt and the absolute volt, using the conversion factor given by Driscoll and Cutkosky[31]

$$1 \text{ NBS volt} = 1.000\,010 \text{ absolute volts} \pm 5 \text{ ppm}.$$

ΔE is the incremental voltage measured by the K3 potentiometer.

As shown in Table I the weighted average of the above measurements gives

$$V\lambda_s = (12\,373.26 \pm 23 \text{ ppm}) \text{ xu-volts}.$$

Combining this with the commonly accepted value of $\Lambda = 1.002\,02 \pm 20$ ppm and $c = 299\,792.5$ km/sec,[32] one obtains

$$h/e = (1.379\,49 \pm 30 \text{ ppm}) \times 10^{-17} \text{ erg} \cdot \text{sec/esu}.$$

The non-x-ray value[32] is

$$h/e = (1.379\,47 \pm 7 \text{ ppm}) \times 10^{-17} \text{ erg} \cdot \text{sec/esu}.$$

The probable errors overlap, and for the first time the x-ray and non-x-ray values of h/e are in good agreement. However, the probable error of the x-ray value is more than four times that of the non-x-ray one. Approximately half of this is due to the uncertainty in the value of Λ.

DuMond and Cohen[33] have pointed out that the optical calibration lines used by Tyrén[34] in his concave grating measurement of Λ require correction for the Lamb shift. This might increase his reported value of $\Lambda = 1.001\,99$ by the order of 50 ppm (due to loss of data and plates it is impossible to determine the exact correction). DuMond and Kirkpatrick in a recent report to the National Science Foundation on a new concave grating experiment give a provisional value

$$\Lambda = 1.001\,80 \pm 10 \text{ ppm}.$$

This is about 200 ppm lower than Tyrén's value or the plane grating values. Thus the true value of Λ may be considerably outside the $1.002\,03 \pm 20$ ppm range calculated by Birge,[35] based principally on measurements with plane gratings.

From these considerations the measured value of $V\lambda_s$ could best be used to evaluate Λ. With $h/e = (1.37947 \pm 7 \text{ ppm}) \times 10^{-17}$ erg·sec/esu[32] the value is

$$\Lambda = 1.002\,005 \pm 24 \text{ ppm}.$$

This value is based on Cu $K\alpha_1 = 1537.400$ xu. However, a report by one of the authors (J. A. B.)[36] on a serious wavelength discrepancy of approximately 20 ppm between the Mo $K\alpha_1$ and the Cu $K\alpha_1$ xu standards has been confirmed and is in preparation for publication. On the basis of Mo $K\alpha_1$ as 707.831 xu (which is taken as a provisional definition of the xu in this report[37]) the wavelength of the Cu $K\alpha_1$ is 1537.370 xu. With this revision the experimental value becomes

$$V\lambda_s = (12\,373.02 \pm 23 \text{ ppm}) \text{ xu} \cdot \text{V}$$

and

$$\Lambda = 1.002\,024 \pm 24 \text{ ppm}.$$

(This revision will not significantly alter the value of h/e calculated above, since the change in Cu $K\alpha_1$ wavelength also implies a compensating change in the ruled grating values of Λ.)

CONCLUSION

The absence of structure and the sharpness of the high-frequency limit of the continuous x-ray spectrum, together with a measurement of the work function correction made possible a precision determination of $V\lambda_s$. The lack of structure in the gas target isochromat tends to support Ulmer's theory as to the origin of the structure. The profile of the isochromat gave the shape of the continuous spectrum produced by single electron bombardment. The measured intensity was constant from a few eV of cutoff to 500 eV above in accordance with the Sommerfeld theory.

With the use of the non-x-ray value of h/e, a precision determination of Λ has been obtained. The accuracy of this value appears comparable to those obtained by either ruled-grating or crystal measurements.

ACKNOWLEDGMENTS

The authors are greatly indebted to Dr. John S. Thomsen for his interest in the work and his assistance in the preparation of the manuscript, to Dr. J. L. Thomas and Dr. W. J. Hamer of the National Bureau of Standards for the recalibration of the electrical standards, and to Dr. U. Fano for theoretical discussions of the shape of the continuous x-ray spectrum at cutoff.

[31] R. L. Driscoll and R. D. Cutkosky, J. Res. Natl. Bur. Std. **60**, 297 (1958).
[32] Natl. Bur. Std. Tech. News Bull. **47**, 175 (1963).
[33] J. W. M. DuMond and E. R. Cohen, Phys. Rev. **103**, 1583 (1956).
[34] Folke Tyrén, dissertation, Uppsala, 1940 (unpublished).
[35] R. T. Birge, Am. J. Phys. **13**, 63 (1945).
[36] J. A. Bearden, Bull. Am. Phys. Soc. **7**, 339 (1962).
[37] J. A. Bearden, A. Henins, J. G. Marzolf, W. C. Sauder, and J. S. Thomsen (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Inelastic Scattering of Electrons by the Hydrogen Molecule Ion*

JAMES M. PEEK
Sandia Laboratory, Albuquerque, New Mexico
(Received 23 December 1963)

The inelastic scattering of electrons by H_2^+ is investigated by means of the first Born approximation. Using the exact electronic wave functions, the integral that must be evaluated to find the differential cross section for fixed internuclear separation is studied for an electronic excitation from the ground state to any discrete excited state. Values related to this integral are tabulated for the processes $1s\sigma_g$–$2p\sigma_u$, $1s\sigma_g$–$2p\pi_u$, $1s\sigma_g$–$2s\sigma_g$ at the equilibrium internuclear distance, while the first case is studied at two additional internuclear separations. The corresponding total cross sections are calculated for incident energies up to 400 eV. Assuming that rotational and vibrational levels of the final electronic state cannot be resolved, it is known that an observed cross section necessarily depends on the initial vibrational state of the molecular ion. The effect of two different initial vibrational states ($\nu=0,3$) is investigated for the $1s\sigma_g$–$2p\sigma_u$ case and it is observed that the vibrational state has a marked influence on the total cross section.

INTRODUCTION

THE problem of inelastic electron scattering by molecules has a complication in addition to those found when considering scattering by atomic systems in that the internal degrees of freedom of the molecule must be taken into consideration. However, in the range of incident energies where the first Born approximation can be expected to be valid, experimentally determined cross sections which resolve rotational effects on electronic transitions have yet to be published, and it appears that the resolution of vibrational structure is just now becoming an experimental possibility. This being the case, it seems highly plausible to treat the internal degrees of freedom so that specific vibrational and rotational excitations are ignored while considering a given electronic process. A detailed analysis of the first Born approximation to scattering by molecules, where the above treatment of the internal modes was employed, has been given by Lassettre[1] and by Craggs and Massey.[2] Following their treatment, the differential cross section, where the excitation is $n\rightarrow n'$ for the electronic system, $L,M\rightarrow L',M'$ for rotation and $\nu\rightarrow\nu'$ for vibration, can be written

$$I(nLM\nu,n'L'M'\nu';\theta)$$
$$=(4k_{n'}/k_nK^4)\left|\int_0^\infty\int_0^{2\pi}\int_0^{\pi}\epsilon(K,\delta,\xi,R)\cdot\Theta_{LM}(\delta,\xi)\Theta_{L'M'}{}^*(\delta,\xi)\mathrm{X}_\nu(R)\mathrm{X}_{\nu'}{}^*(R)R^2\times\sin\delta d\delta d\xi dR\right|^2. \quad (1)$$

Θ, X are rotation and vibration functions, respectively, where it has been assumed that these functions do not mutually interact in addition to the usual assumption that the nuclear and electronic coordinates are separable. This expression and all remaining equations are given in atomic units unless stated otherwise. The quantity ϵ is defined for a molecule with j electrons as (neglecting spin variables)

$$\epsilon(K,\delta,\xi,R)=\int\Psi_{n'}{}^*(\mathbf{r}_1\cdots\mathbf{r}_j;R)\left[\sum_{i=1}^{j}\exp(i\mathbf{K}\cdot\mathbf{r}_i)\right]\times\Psi_n(\mathbf{r}_1\cdots\mathbf{r}_j;R)d\mathbf{r}_1\cdots d\mathbf{r}_j, \quad (2)$$

where the states n,n' cannot be the same. The quantities K, δ, ξ, R refer to the parameters that must be fixed during the integration over the electronic coordinates. The symbol K, the magnitude of $\mathbf{K}$, defines the momentum change of an incident electron scattered through an angle θ; $K^2=k_n{}^2+k_{n'}{}^2-2k_nk_{n'}\cos\theta$, where k_n, $k_{n'}$ are the magnitudes of the momentum for the incident electron before and after scattering. The angles δ, ξ fix the orientation of the molecule in space and R is the internuclear distance. Lassettre[1] has shown that the differential cross section becomes

$$I(n\nu,n'\nu';\theta)=(4k_{n'}/k_nK^4)\left\{(1/4\pi)\times\int_0^{2\pi}\int_0^{\pi}\left|\int_0^\infty\epsilon(K,\delta,\xi,R)\mathrm{X}_\nu(R)\mathrm{X}_{\nu'}{}^*(R)R^2dR\right|^2\times\sin\delta d\delta d\xi\right\} \quad (3)$$

when one sums over all final rotational states and averages over the degenerate levels of the initial state. Equation (3) will be true if the rotation of the molecule can be described by a symmetrical top and if the dependence of $k_{n'}$, K on $L'M'$ is ignored.[1] It has been

* This work performed under the auspices of the U. S. Atomic Energy Commission.

[1] E. N. Lassettre, Ohio State University Research Foundation, Columbus, Ohio, 1957, R. F. Project 464, Report No. 1 (unpublished).

[2] J. D. Craggs and H. S. W. Massey, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1959), Vol. 37, Part 1, p. 332.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

assumed that the final vibrational levels cannot be resolved, so a sum over all ν' can be made, which gives

$$I(n\nu,n';\theta)=(4k_{n'}/k_nK^4)\Big\{(1/4\pi)\times\int_0^\infty\int_0^{2\pi}\int_0^{\pi}|\epsilon(K,\delta,\xi)|^2|X_\nu(R)|^2R^2\sin\delta\times d\delta d\xi dR\Big\} \quad (4)$$

if the dependence of $k_{n'}$, K on ν' is neglected. This is then the first Born approximation for a given electronic excitation of a molecule when the final rotational and vibrational states of the system are unresolved.

The simplest molecular system available for study is the hydrogen molecule ion. Previous calculations on electron scattering by this system have been reported by Ivash[3] and Kerner.[4] Ivash treats the electronic excitation $1s\sigma_g$–$2p\sigma_u$ (in the notation of Bates, Ledsham and Stewart[5]), where he approximates to the ground vibrational state by fixing $R=R_e$, the equilibrium internuclear distance, and then evaluates Eq. (4) by ignoring the integration over R. In evaluating Eq. (2), he also treats the possibility of electron exchange. Kerner does not use Eq. (4), but evaluates Eq. (1) by considering an excitation to a specific rotational and vibrational state while treating the $1s\sigma_g$–$2p\sigma_u$ electronic transition. In each of these studies, the electronic wave functions that appear in Eq. (2) were replaced by linear combinations of $1s$ atomic orbitals. Since the exact wave functions for H_2^+ (assuming fixed nuclei) are available[5] and the use of exact wave functions for the scatterer is assumed in the derivation of the first Born approximation, it would seem that electron scattering by H_2^+ should be reexamined using these wave functions. This will make it possible to give a quantitative estimate of the error introduced into the first Born approximation when approximate electronic wave functions are used to evaluate Eq. (2), at least for this relatively simple case. In addition, there are a number of electronic processes which can be expected to contribute to the inelastic scattering by H_2^+ and it would be interesting to obtain quantitative information about some of these electronic transitions.

The integral defined in Eq. (2) is studied for transitions from the ground state to any discrete excited electronic state. It is found that this integral can be evaluated analytically with respect to all coordinates except one. The behavior of this integral as a function of K and the orientation parameters is explored and a quantitative example is given for one transition. A table of this integral, after the appropriate integration over nuclear orientation has been performed, is given for the transitions $1s\sigma_g$–$2p\sigma_u$, $1s\sigma_g$–$2p\pi_u$, $1s\sigma_g$–$2s\sigma_g$. All transitions are studied at the equilibrium internuclear distance while the $1s\sigma_g$–$2p\sigma_u$ case is evaluated at two additional internuclear separations. The total cross section for incident energies up to 400 eV and fixed R is then calculated for these various cases. The effect of two different vibrational states for the ground state, $\nu=0$, 3, is also studied for the $1s\sigma_g$–$2p\sigma_u$ transition and the corresponding total cross sections are given.

GENERAL CONSIDERATIONS

The total cross section corresponding to Eq. (4) is

$$Q_\nu=\int I(n\nu,n';\theta)d\Omega=(8\pi/k_n^2)(1/4\pi)\times\int_{kn-kn'}^{kn+kn'}\int_0^\infty\int_0^{2\pi}\int_0^{\pi}\left\{\frac{|\epsilon(K,\delta,\xi,R)|^2}{K^3}\right\}|X_\nu(R)|^2\times R^2\sin\delta d\delta d\delta d\xi dRdK, \quad (5)$$

where $d\Omega$ is the solid-angle volume element. Carrying out the integration over the orientation angles,[1]

$$|\epsilon(K,R)|^2=(1/4\pi)\int_0^{2\pi}\int_0^{\pi}|\epsilon(K,\delta,\xi,R)|^2\sin\delta d\delta d\xi \quad (6)$$

is found. Defining

$$Q(R)=(8\pi/k_n^2)\int_{kn-kn'}^{kn+kn'}[|\epsilon(K,R)|^2/K^3]dK \quad (7)$$

and interchanging the order of integration, the total cross section becomes

$$Q_\nu=\int_0^\infty Q(R)|X_\nu(R)|^2R^2dR. \quad (8)$$

In order to obtain $Q(R)$ the integrations defined by Eqs. (2), (6), (7) must be carried out for a fixed R. Then $Q(R)$ must be found as a function of R to effect the evaluation of Eq. (8).

For the hydrogen molecule ion, Eq. (2) becomes

$$\epsilon(K,\delta,R)=\int_1^\infty\int_0^{2\pi}\int_{-1}^{1}\exp\{(iKR/2)[\lambda\mu\cos\delta+[(\lambda^2-1)(1-\mu^2)]^{1/2}\sin\phi\sin\delta]\}\times\Psi_n\Psi_{n'}^*(R/2)^3(\lambda^2-\mu^2)d\mu d\phi d\lambda, \quad (9)$$

where the coordinate system is shown in Fig. 1. It should be noted that for any linear scattering system ξ can be suppressed. The wave functions to be used in evaluating Eq. (9) are given by Bates, Ledsham, and

[3] E. V. Ivash, Phys. Rev. 112, 155 (1958).
[4] E. H. Kerner, Phys. Rev. 92, 1441 (1953).
[5] D. R. Bates, Kathleen Ledsham, and A. L. Stewart, Phil. Trans. Roy. Soc. London A246, 215 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

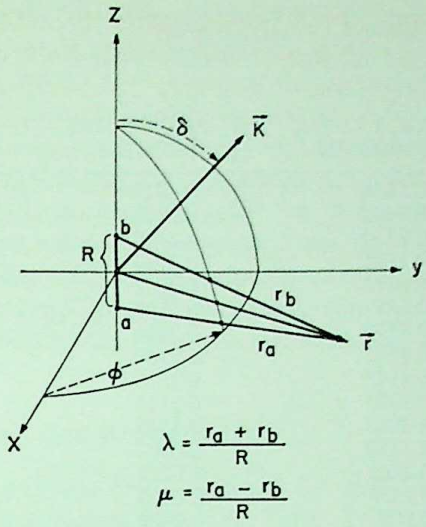


FIG. 1. The coordinate system for H_2^+ and its orientation with respect to K.

Stewart[5] and are of the form

$$\Psi_n = \left[(\lambda+1)^{\sigma}(\lambda^2-1)^{m/2} e^{-p\lambda} \sum_i q_i \left(\frac{\lambda-1}{\lambda+1}\right)^i \right] \times \left[\sum_t{}' f_t P_{t+m}{}^m(\mu)\right] e^{\pm im\phi}. \quad (10)$$

The primed sum means that only even or odd values of t are to be used. The quantities σ, p, m are eigenparameters for the nth state and $P_{t+m}{}^m$ is an associated Legendre polynomial of the first kind. The product $\Psi_n \Psi_{n'}{}^*$ appearing in Eq. (9) can be written in the form

$$\mu^2 \Psi_n \Psi_{n'}{}^* = \mu^2 f(\lambda) \exp(im\phi) \sum_t{}' a_t P_{t+m}{}^m(\mu)$$

$$= f(\lambda) \exp(im\phi) \sum_t{}' b_t P_{t+m}{}^m(\mu), \quad (11)$$

since the product of any two Legendre polynomials can be written as a sum of Legendre polynomials. [See Ref. 6 for a general formula applicable to the case $m=0$.] The ground state involves only terms of the type P_{2k}; hence, the sum in Eq. (11) will have values of m and $t+m$ identical to those representing the excited state.

In this notation Eq. (9) becomes

$$\epsilon(K,\delta,R) = 4\pi(R/2)^3 \left\{ \sum_n{}' i^n a_n \times \int_1^{\infty} \lambda^2 f(\lambda) j_{n+m}(y) P_{n+m}{}^m(\cos\psi) d\lambda - \sum_n{}' i^n b_n \int_1^{\infty} f(\lambda) j_{n+m}(y) P_{n+m}{}^m(\cos\psi) d\lambda \right\}, \quad (12)$$

where

$$y = (KR/2)(\lambda^2 - \sin^2\delta)^{1/2}, \quad (13)$$

$$\cos\psi = \lambda \cos\delta / (\lambda^2 - \sin^2\delta)^{1/2}, \quad (14)$$

and j is a spherical Bessel function. The details of this derivation are sketched in the Appendix.

NUCLEAR ORIENTATION

The qualitative behavior of $\epsilon(K,\delta,R)$ as a function of δ for various final states can easily be determined from Eq. (9) or Eqs. (12)–(14). Expanding the exponential of Eq. (9) in a power series of the exponent and then noticing the symmetry of the integral with respect to μ or ϕ, a large number of terms will be found to vanish. In addition, each nonvanishing term in this series will have a very restricted dependence on δ. Listing the results for the cases studied here, it is found all nonvanishing terms involve $\cos(2k+1)\delta$ for final $np\sigma_u$ states, $\sin(2k+1)\delta$ for final $np\pi_u$ states, and $\cos 2k\delta$ for final $ns\sigma_g$ states, where n, k are positive integers. These same conclusions can be reached by using Eqs. (12)–(14). It is also known that $\epsilon(K,\delta,R)$ must be either symmetric or antisymmetric about $\delta=0$, $\pi/2$, π, since $|\epsilon(K,\delta,R)|^2$ must be symmetric about these points. This follows immediately from the symmetry of H_2^+ and implies that we need $\epsilon(K,\delta,R)$ only over the range $0 \leq \delta \leq \pi/2$.

From the above analysis or directly from Eqs. (12)–(14), it can be shown that for certain transitions $\epsilon(K,\delta,R)$ will vanish identically for a specific orientation. It is apparent that $\epsilon(K,\delta,R)$ is zero when $\delta=\pi/2$ for $np\sigma_u$ final states and when $\delta=0$ for $np\pi_u$ final states. $\epsilon(K,\delta,R)$ is never identically zero for final states of the type $ns\sigma_g$. This type of orientation selection rule has been noticed by other authors and the results here are consistent with the conclusions of Dunn[7] for the homonuclear diatomic molecule.

The quantitative behavior of $\epsilon(K,\delta,R)$ as a function of K and δ is shown in Fig. 2 for the case $1s\sigma_g$–$2p\sigma_u$, $R=2.0$. In the limit of small K, $\epsilon(K,\delta,R)$ approaches

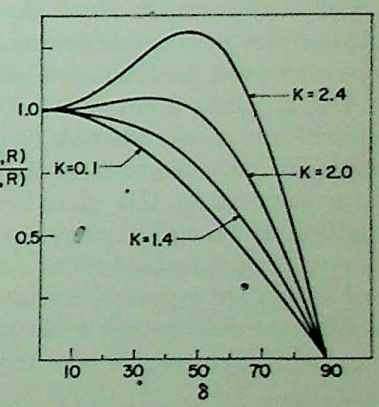


FIG. 2. The behavior of $\epsilon(K,\delta,R)$ shown as a function of δ for the $1s\sigma_g$–$2p\sigma_u$ ($R=2.0$) case and several values of K.

[6] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, 1958), p. 331.

[7] G. H. Dunn, Phys. Rev. Letters 8, 62 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Values of $|\epsilon(K,R)|^2/K^2$ [see Eq. (6)] for the electronic transitions $1s\sigma_g$–$2p\sigma_u$ ($R=1.4$, 2.0, 3.2), $1s\sigma_g$–$2p\pi_u$ ($R=2.0$), $1s\sigma_g$–$2s\sigma_g$ ($R=2.0$) as a function of K.

	$\lvert\epsilon(K,R)\rvert^2/K^2$				
K	$1s\sigma_g$–$2p\sigma_u$ ($R=2.0$) ($\Delta E=0.43509$)	$1s\sigma_g$–$2p\sigma_u$ ($R=1.4$) ($\Delta E=0.67219$)	$1s\sigma_g$–$2p\sigma_u$ ($R=3.2$) ($\Delta E=0.18102$)	$1s\sigma_g$–$2p\pi_u$ ($R=2.0$) ($\Delta E=0.67385$)	$1s\sigma_g$–$2s\sigma_g$ ($R=2.0$) ($\Delta E=0.74176$)
0.0	0.3676	0.2296	0.7656	0.3414	0.0
0.05	0.3666	0.2291	0.7632	0.3403	
0.1	0.3641	0.2277	0.7559	0.3369	0.001091
0.2	0.3539	0.2221	0.7275	0.3237	0.004143
0.3	0.3378	0.2130	0.6826	0.3031	0.008633
0.4	0.3166	0.2011	0.6247	0.2765	0.01382
0.5	0.2915	0.1869	0.5579	0.2461	0.01890
0.6	0.2638	0.1711	0.4867	0.2140	0.02320
0.7	0.2349	0.1543	0.4145	0.1818	0.02622
0.8	0.2058	0.1373	0.3453	0.1513	0.02777
0.9	0.1777	0.1205	0.2816	0.1233	0.02785
1.0	0.1512	0.1045	0.2250	0.0987	0.02667
1.1	0.1270	0.0896	0.1762	0.0777	0.02456
1.2	0.1053	0.0760	0.1355	0.0602	0.02185
1.3	0.0863	0.0639	0.1024	0.0460	0.01886
1.4	0.0700	0.0532	0.0762	0.0347	0.01585
1.5	0.0562	0.0439	0.0558	0.0259	0.01301
1.6	0.0447	0.0360	0.0404	0.0191	0.01047
1.7	0.0353	0.0293	0.0289	0.0140	0.00827
1.8	0.0277	0.0237	0.0205	0.0102	0.00643
1.9	0.0215	0.0191	0.0145	0.0073	0.00493
2.0	0.0166	0.0153		0.0053	0.00374
2.1					0.00281
2.2	0.0098				

cosine behavior; hence, the differential cross section will have a cosine-squared dependence on δ. As K increases, the behavior changes from a cosine curve to one quite eccentric, although still possessing cosine symmetry. Note that the maximum in $\epsilon(K,\delta,R)$ moves away from $\delta=0$ and approaches $\delta=\pi/2$. From the form of Eq. (9), when the exponential is expanded in a power series, it is possible to say something about the behavior of this eccentricity as the internuclear separation is altered. As $R\rightarrow 0$, $\epsilon(K,\delta,R)$ becomes much less dependent on K while, as R is allowed to become large, its behavior becomes much more dependent on K and quite complex in general. This behavior is observed in the case where several internuclear separations are studied. These general arguments apply to other transitions but both the symmetry and quantitative behavior can be expected to be different.

These observations about the dependence of $\epsilon(K,\delta,R)$ are useful when the integration of Eq. (6) is performed. $\epsilon(K,\delta,R)$ need be evaluated only over the range $0\leq\delta\leq\pi/2$, as pointed out above. Since this must be done numerically, a considerable saving in labor ensues. In addition, since the form of the trigonometric behavior is well specified, this gives a strong hint as to the method of numerical integration that should be used. If one evaluates $\epsilon(K,\delta,R)$ at a series of equally spaced values of δ for a given K and then fits these points with a trigonometric interpolation formula,[8] a simple series results which contains only terms like $\cos(2n+1)\delta$ for the $1s\sigma_g$–$2p\sigma_u$ case. Similar behavior results for the other cases. It is a simple matter to then square this series and carry out the indicated integration. In practice, $\epsilon(K,\delta,R)$ was evaluated for increments of 10 deg in δ; hence, the series consists of nine terms. This procedure is exact if the expansion of the exponential in Eq. (9) can be truncated to the first nine nonvanishing terms. Obviously, as K and R increase, this expansion becomes less valid. However, for the cases studied, a sufficient range of K could be covered so that all significant contributions to the total cross section were obtained. The resulting values of $|\epsilon(K,R)|^2/K^2$ for the various cases are given in Table I. An accurate error analysis seems out of the question but it is estimated that any error should be confined to the last figure quoted.

One check of the values given in Table I can be made through the relationship

$$\lim_{K\rightarrow 0} f(K,R)\rightarrow f(R), \tag{15}$$

where

$$f(K,R)=2\Delta E\frac{|\epsilon(K,R)|^2}{K^2}. \tag{16}$$

$f(R)$ is the optical oscillator strength and $f(K,R)$ is the generalized optical oscillator strength as defined by Bethe.[9] Values of $f(R)$ for the $1s\sigma_g$–$2p\sigma_u$ case are

[8] C. Lanczos, *Applied Analysis* (Prentice Hall, Inc., Englewood Cliffs, N. J., 1956), p. 229.

[9] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Oxford University Press, Oxford, 1952), p. 248.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

available[10] and are 0.309 ($R=1.4$), 0.319 ($R=2.0$), 0.281 (R=3.2) which are to be compared with 0.3087, 0.3198, 0.2772, the values of $f(R)$ obtained from Eq. (15). The optical oscillator strength for the $1s\sigma_g$–$2p\pi_u$ transition[11] is 0.460 which is to be compared to 0.4601. Obviously, $f(R)$ vanishes for the forbidden $1s\sigma_g$–$2s\sigma_g$ transition, as does the limit of Eq. (15) for this case. [Nonvanishing limits were found directly from Eq. (9) rather than using Eq. (12).] Agreement is quite good for all cases except $1s\sigma_g$–$2p\sigma_u$ ($R=3.2$) and the value obtained here is within the limits of accuracy (3%) claimed by Bates.[10]

TOTAL CROSS SECTIONS

The total cross sections $Q(R)$, defined by Eq. (7), were found from the values listed in Table I in the following way: For small values of K ($K\leq 0.4$) Table I was interpolated graphically to intervals in K of 0.01. Then the integrand of Eq. (7) was evaluated for available values of $|\epsilon(K,R)|^2/K^2$ and the integration was carried out using Simpson's rule. In general, Table I does not contain values of $|\epsilon(K,R)|^2/K^2$ for K large enough to include the upper limit of the integral in Eq. (7). However, the contributions to the total cross section from this range are quite small and in practice this remainder was estimated by fitting the integrand with a function depending on K^{-5}.

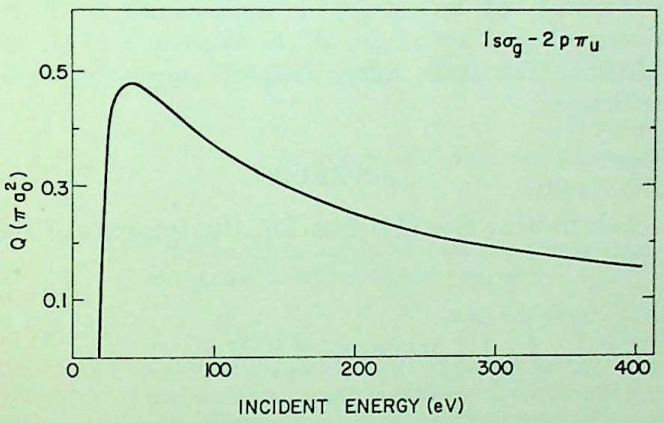


FIG. 3. The total cross section $Q(R)$ for the $1s\sigma_g$–$2p\pi_u$ ($R=2.0$) case.

The cross section $Q(2.0)$ for the $1s\sigma_g$–$2p\pi_u$ case is shown in Fig. 3. (The equilibrium internuclear distance of H_2^+ is $R_e=2.0\ a_0$.) It is apparent that this transition contributes quite heavily to scattering by H_2^+ and must be taken into account when considering processes where this final state is possible. The $1s\sigma_g$–$2s\sigma_g$, $R=2.0$ case is shown in Fig. 4. This corresponds to an optically forbidden transition and the total cross section is relatively small, as expected. The $1s\sigma_g$–$2p\sigma_u$ case is shown in Fig. 5 for three internuclear separations. The curve for $R=2.0$ corresponds to the equilibrium internuclear separation and is, therefore, an approximation to the Born cross section when the effects of the integration in Eq. (8) are ignored. This is, of course, most accurate when H_2^+ is in the $v=0$ vibrational state. The other two curves are for internuclear separations that correspond roughly to the turning points of the vibrational state $v=3$. It is apparent that the cross section for this case is strongly dependent on internuclear distance for intermediate incident energies. In the limit of infinite incident energies, these cross sections approach constant ratios of each other; specifically, $Q(3.2)/Q(2.0)=2.1$ and $Q(2.0)/Q(1.4)=1.6$. The effect of this strong dependence of $Q(R)$ on R is investigated below where Eq. (8) is evaluated for two different vibrational states.

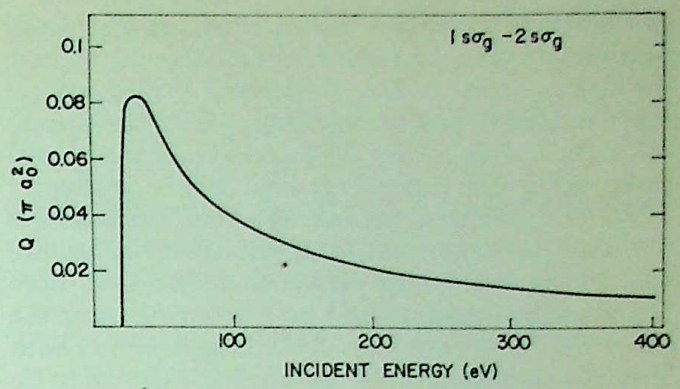


FIG. 4. The total cross section $Q(R)$ for the $1s\sigma_g$–$2s\sigma_g$ ($R=2.0$) case.

The curve for $1s\sigma_g$–$2p\sigma_u$ ($R=2.0$) can be compared with the results, ignoring exchange, given by Ivash[3]

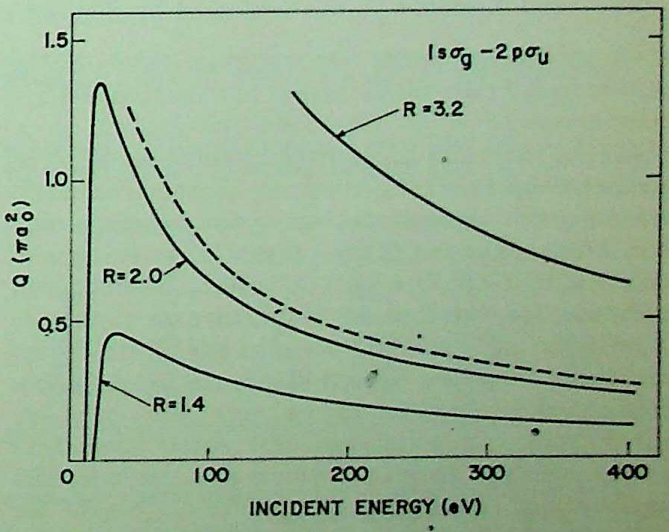


FIG. 5. The total cross section $Q(R)$ for the $1s\sigma_g$–$2p\sigma_u$ ($R=1.4$, 2.0, 3.2) case. The dashed line is from a calculation by Ivash (Refs. 3, 12).

[10] D. R. Bates, J. Chem. Phys. 19, 1122 (1951).
[11] D. R. Bates, R. T. S. Darling, S. C. Hawe, and A. L. Stewart, Proc. Phys. Soc. (London) A66, 1124 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which are shown as the dashed curve[12] in Fig. 5. To obtain this curve, he used wave functions of the form $\Psi=u_a+u_b$, where u is a $1s$ screened hydrogen wave function. His results are uniformly above the present calculation of $Q(2.0)$ by 20% or more in this energy range and in the high-energy limit will be greater by 17%. The difference between the two calculations is not large but the present treatment has the advantage that all aspects of the derivation of the first Born approximation remain valid as a result of using the exact H_2^+ wave functions. A direct comparison with Kerner's calculation does not seem justified since he treats a specific rotational and vibrational excitation in conjunction with the $1s\sigma_g$–$2p\sigma_u$ electronic transition. However, it is interesting to note that his cross section is roughly an order of magnitude less than $Q(2.0)$.

To study the effect of vibration on the total cross section, it is necessary to evaluate Eq. (8). If a diatomic molecule is in its lowest vibrational state ($\nu=0$) and it is assumed that $Q(R)$ varies slowly with R, it is a good approximation[2] to replace Q_0 with $Q(R_e)$, since $|X_0(R)|^2$ will be strongly peaked at the equilibrium internuclear distance R_e. For the transition under consideration, $Q(R)$ is seen to be rather strongly dependent on R, so it appears that Q_0 must be evaluated without using the above approximation. In addition, it is known that H_2^+ is often observed in excited vibrational states which have a rather long lifetime. Hence, Q_ν will be evaluated using Eq. (8) for the vibrational state $\nu=3$. This state is chosen because it is a highly probable vibrational state for H_2^+ if this ion results from the ionization of H_2 in its ground electronic and vibrational state.

The integral defined by Eq. (8) was estimated in the following way: $Q(R)$ for a given incident energy was fit by a polynomial in R. Since $Q(R)$ is known for only three points this polynomial was necessarily a quadratic. Then, using the vibrational functions given by Cohen, Hiskes, and Riddell,[13] Q_ν was evaluated by Simpson's rule.

The results for Q_0, Q_3 are shown in Table II along with $Q(1.4)$, $Q(2.0)$, $Q(3.2)$. It is interesting to note that Q_0 is about 10% higher than $Q(2.0)$ and that Q_3 is almost 70% higher at 400 eV with the deviation increasing as the incident energy decreases. Hence, to predict the inelastic electron scattering into the $2p\sigma_u$ state, it is necessary to know precisely what vibrational levels are populated, and to what degree. It does not follow that scattering into other final electronic states will show such marked dependence on ν, but points out the necessity for caution before neglecting the initial vibrational state. Relationships equivalent to Eq. (8) can easily be derived for the differential cross section and the generalized oscillator strength. The effects of ν on the differential and total cross sections can be expected to be quite similar. However, using the information in Table I, it can be seen that $f(K,R)$ has a maximum near R_e for the $1s\sigma_g$–$2p\sigma_u$ case; hence, the generalized oscillator strength will be much less dependent on ν.

TABLE II. The total cross sections $Q(R)$ and Q_ν (in units of πa_0^2) for the $1s\sigma_g$–$2p\sigma_u$ transition tabulated as a function of the incident energy.

E(eV)	$Q(1.4)$	$Q(2.0)$	$Q(3.2)$	Q_0	Q_3
50	0.426	0.977	3.25	1.09	1.92
100	0.304	0.635	1.92	0.698	1.17
200	0.197	0.396	1.11	0.431	0.691
300	0.148	0.295	0.795	0.320	0.502
400	0.120	0.230	0.612	0.248	0.383

[12] Ivash's calculation has been extended to include the energy range of interest in this investigation.

[13] S. Cohen, J. R. Hiskes, and R. J. Riddell, Jr., University of California Radiation Laboratory Report No. UCRL-8871, 1959 (unpublished); S. Cohen, J. R. Hiskes, and R. J. Riddell, Jr., Phys. Rev. 119, 1025 (1960).

ACKNOWLEDGMENTS

I would like to express my appreciation to T. A. Green, C. J. MacCallum, W. E. Warren, and G. W. McClure for their many helpful suggestions and criticisms.

APPENDIX

Substituting Eq. (11) into Eq. (9), integrals of the type

$$I_k^m=\int_1^\infty\int_{-1}^1\int_0^{2\pi} f(\lambda)\exp\{(iKR/2)\times[\lambda\mu\cos\delta+(\lambda^2-1)^{1/2}(1-\mu)^{1/2}\sin\phi\sin\delta]-im\phi\}P_{k+m}^m(\mu)d\phi d\mu d\lambda \quad \text{(A1)}$$

are obtained. Utilizing the integral[14]

$$\int_\alpha^{2\pi+\alpha} e^{i(n\phi-Z\sin\phi)}d\phi=2\pi J_n(Z), \quad \text{(A2)}$$

where J_n is a Bessel function of the first kind,

$$I_k^m=2\pi\int_1^\infty\int_{-1}^1 f(\lambda)\exp\{i(KR/2)\lambda\mu\cos\delta)\}\times J_m[\tfrac{1}{2}KR(\lambda^2-1)^{1/2}(1-\mu^2)^{1/2}\sin\delta]\times P_{k+m}^m(\mu)d\mu d\lambda \quad \text{(A3)}$$

is found. Gegenbauer's finite integral[15] can be rewritten

[14] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, Cambridge, 1958), p. 20.

[15] Ref. 14, p. 378.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in the form

$$\int_{-1}^{1} \exp(iy\mu\cos\psi) J_m[y(1-\mu^2)^{1/2}\sin\psi] P_{k+m}{}^m(\mu) d\mu = 2i^k P_{k+m}{}^m(\cos\psi)\, j_{k+m}(y), \quad \text{(A4)}$$

where $j_n(x)$ is a spherical Bessel function defined by

$$j_n(x) = (\pi/2x)^{1/2} J_{n+1/2}(x). \quad \text{(A5)}$$

Let

$$y = (KR/2)(\lambda^2 - \sin^2\delta)^{1/2} \quad \text{(A6)}$$

$$\cos\psi = \lambda\cos\delta/(\lambda^2-\sin^2\delta)^{1/2}; \quad \text{(A7)}$$

then

$$\sin\psi = (\lambda^2-1)^{1/2}\sin\delta/[\lambda^2-\sin^2\delta]^{1/2}. \quad \text{(A8)}$$

Substituting Eqs. (A6)–(A8) into Eq. (A4), we find an integral of the same form that appears in Eq. (A3), and it follows that

$$I_k{}^m = 4\pi i^k \int_1^\infty f(\lambda) j_{k+m}(y) P_{k+m}{}^m(\cos\psi) d\lambda. \quad \text{(A9)}$$

Using these results, Eqs. (12)–(14) are easily verified.

Assuming that Eq. (A9) cannot be evaluated analytically, the following numerical scheme was used for this purpose. A program for a digital computer, the Control Data Corporation's 1604, was written which carried out the integration of Eq. (A9) for a given choice of K, δ, R by means of Weddle's quadrature formula. With these values it is then possible to evaluate Eq. (12). The sum in Eq. (12) was usually truncated to three terms, although four terms were used in some cases. The same program was used to evaluate $\epsilon(K,\delta,R)$ and carry out the integration of Eq. (6).

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

$\Delta S = \pm 1$ Magnetic Multipole Radiative Transitions*

MASATAKA MIZUSHIMA
Department of Physics and Astrophysics, University of Colorado, Boulder, Colorado

(Received 20 December 1963)

The magnetic multipole transition probability is calculated in terms of the matrix elements of the magnetic multipole. The magnetic jm moment $Q_{jm}{}^{(mg)}$ is defined as

$$(e/\mu)[4\pi/(2j+1)]^{1/2} \sum_i (\nabla r_i{}^j Y_{jm})[(j+1)^{-1}\mathbf{l}_i + \mathbf{s}_i],$$

where e and μ are electron charge and electron mass, r_i, $\mathbf{l}_i$, and $\mathbf{s}_i$ are the coordinate, orbital angular momentum, and spin-angular momentum of the ith electron and Y_{jm} is the spherical harmonic. Magnetic quadrupole and octupole moments are explicitly given. It is shown that for the ${}^3\Sigma_u{}^+ \leftrightarrow {}^1\Sigma_g{}^+$ transition of the hydrogen molecule, the magnetic quadrupole transition is more important than the conventional spin-orbit electric dipole transition. The magnetic octupole transition has the same order of magnitude as the spin-orbit magnetic dipole transition.

INTRODUCTION

THE transition between states with different multiplicity is a very weak one. The corresponding emission line of some ions were first found in the spectra of some nebulas[1] and the theory has been given by Condon and other people.[2] This type of transition has been investigated for many ions since then, a brief review can be found in Garstang's article.[3] In all these theories the mixing of different multiplicity states through the spin-orbit interaction was assumed to be important. The ${}^1D_2$ state of O III, for example, has 0.0074 of ${}^3P_2$ wave function mixed through the spin-orbit interaction, and the transition to a triplet state can occur through this small part of the wave function. Most of the phosphorescences of molecules are interpreted as due to the singlet-triplet transition.[4] McClure first suggested the transition through the spin-orbit interaction for these molecules.[5] The present author and Koide formed a theory starting with the Dirac equation of electron and calculated the singlet-triplet transition probability of the benzene molecule.[6] The result with some approximation gave about 1 sec for the lifetime of the lowest ${}^3B_{2u}$ state of this molecule. Hameka and Oosterhoff[7] refined the calculation and showed the lifetime of the same state to be 190 sec. Calculations for

* Supported by National Science Foundation.

[1] I. S. Bowen, Astrophys. J. 67, 1 (1928).
[2] See C. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, New York, 1935), pp. 282–283.
[3] R. H. Garstang, *Atomic and Molecular Processes*, edited by D. R. Bates (Academic Press Inc., New York, 1962), Part 1.
[4] M. Kascha, Chem. Rev. 41, 401 (1947).
[5] D. C. McClure, J. Chem. Phys. 17, 665 (1949).
[6] M. Mizushima and S. Koide, J. Chem. Phys. 20, 765 (1952).
[7] H. F. Hameka and L. J. Oosterhoff, Mol. Phys. 1, 358 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

other molecules have also been done following the same idea, and they are briefly reviewed by Garstang.[3]

In this paper a completely different mechanism, namely, the magnetic multipole transition is proposed. If the probability of the electric dipole transition is $P_1^{(el)}$, the conventional theory of $\Delta S=\pm 1$ transitions gives the probability of about $(\hbar\omega/\mu c^2)^2 P_1^{(el)}$ for the case of spin-orbit electric dipole transition, since the spin-orbit coupling is a first-order relativistic correction to the Coulomb energy which is about $\hbar\omega$. $\hbar\omega$ is the photon energy which appears or disappears in our transition and μc^2 is the rest-mass energy of the electron. It will be shown that the magnetic quadrupole transition has the probability of the same order of magnitude, namely, $(\hbar\omega/\mu c^2)P_1^{(el)}$. In the same way one can show that the spin-orbit magnetic dipole transition has the same order of magnitude as the magnetic octupole transition.

MAGNETIC MULTIPOLE TRANSITION PROBABILITY

The theory of the spherical photon is given in Akhiezer and Berestetsky's book.[8] They showed that the electric field for a photon which is the eigenstate of the total angular momentum with eigenvalues j and m, and the eigenstate of the energy with eigenvalue $\hbar\omega$ and at the same time the eigenstate of the parity operator with eigenvalue $(-1)^{j+1}$, is

$$\mathbf{E}_{\omega jm}=\boldsymbol{\mathcal{E}}_{\omega jm}+\boldsymbol{\mathcal{E}}_{\omega jm}^*, \tag{1}$$

$$\boldsymbol{\mathcal{E}}_{\omega jm}=(\hbar\omega^3/16\pi^3\epsilon_0 c^2)^{1/2}e^{-i\omega t}g_j(\omega r/c)\mathbf{Y}_{jjm}(\mathbf{r}/r), \tag{2}$$

where ϵ_0 is the permittivity of the vacuum and

$$g_j(\omega r/c)=(2\pi)^{3/2}i^j J_{j+1/2}(\omega r/c)/(\omega r/c)^{1/2}, \tag{3}$$

with the Bessel function $J_{j+1/2}$, and $\mathbf{Y}_{jjm}(\mathbf{r}/r)$ is the spherical vector given by the spherical harmonics $Y_{jm}(\mathbf{r}/r)$ as

$$\mathbf{Y}_{jjm}(\mathbf{r}/r)=\begin{Bmatrix}[(j+m)(j-m+1)]^{1/2}Y_{jm-1}+[(j-m)(j+m+1)]^{1/2}Y_{jm+1}\\ i[(j+m)(j-m+1)]^{1/2}Y_{jm-1}-i[(j-m)(j+m+1)]^{1/2}Y_{jm+1}\\ 2mY_{jm}\end{Bmatrix}[4j(j+1)]^{-1/2}. \tag{4}$$

There exists another state with the same ω, j, and m eigenvalues, but with the parity of $(-1)^j$. Such state is called an electric state while our state is called a magnetic state. In this paper we are interested in the magnetic states only.

If we take the Coulomb gauge we see from (2) that the vector potential $\mathbf{A}_{\omega jm}$ is

$$\mathbf{A}_{\omega jm}=\mathcal{A}_{\omega jm}+\mathcal{A}_{\omega jm}^*, \tag{5}$$

$$\mathcal{A}_{\omega jm}=i(\hbar\omega/16\pi^3\epsilon_0 c^2)^{1/2}e^{-i\omega t}g_j(\omega r/c)\mathbf{Y}_{jjm}(\mathbf{r}/r). \tag{6}$$

Note that the normalization is such that

$$\epsilon_0\int \boldsymbol{\mathcal{E}}_{\omega jm}\cdot\boldsymbol{\mathcal{E}}_{\omega' j'm'}^* d\mathbf{r}=\tfrac{1}{2}\hbar\omega c\delta(\omega-\omega')\delta_{jj'}\delta_{mm'},$$

$$\int \boldsymbol{\mathcal{E}}_{\omega jm}^*\cdot\boldsymbol{\mathcal{E}}_{\omega' j'm'}^* d\mathbf{r}=0=\int \boldsymbol{\mathcal{E}}_{\omega jm}\cdot\boldsymbol{\mathcal{E}}_{\omega' j'm'} d\mathbf{r}. \tag{7}$$

Thus, if we expand a given field as

$$\mathbf{E}=\sum_{jm}\int F(\omega jm)\boldsymbol{\mathcal{E}}_{\omega jm}d\omega/c^{1/2}+\sum_{jm}\int F^*(\omega jm)\boldsymbol{\mathcal{E}}_{\omega jm}^* d\omega/c^{1/2}, \tag{8}$$

we see the field energy U is

$$U=\epsilon_0\int E^2 d\mathbf{r}=2\epsilon_0\sum_{jm}\sum_{j'm'}\int F(\omega jm)\times F^*(\omega' j'm')\boldsymbol{\mathcal{E}}_{\omega jm}\cdot\boldsymbol{\mathcal{E}}_{\omega' j'm'}^* d\omega d\omega' d\mathbf{r}/c$$
$$=\sum_{jm}\int \hbar\omega F(\omega jm)F^*(\omega jm)d\omega, \tag{9}$$

which indicates that $F(\omega jm)$ and $F^*(\omega jm)$ are annihilation and creation operators for the (ωjm), photon respectively,

$$(\nu_{\omega jm}|F^*(\omega jm)|\nu_{\omega jm}-1)=(\nu_{\omega jm}/d\omega)^{1/2}=(\nu_{\omega jm}-1|F(\omega jm)|\nu_{\omega jm}), \tag{10}$$

where $\nu_{\omega jm}$ is the number of the (ωjm) photon.

The probability of transition $n_1\nu_1\leftrightarrow n_2\nu_2$ is[9]

$$P(n_1\nu_1\leftrightarrow n_2\nu_2)=(\pi/\hbar^2)|(n_1\nu_1|H_i|n_2\nu_2)|^2\delta(\omega_{\nu_1\nu_2}-\omega_{n_1 n_2}), \tag{11}$$

where n_1 and n_2 are molecular states, ν_1 and ν_2 are photon numbers,

$$\omega_{ab}=(E_a-E_b)/\hbar \tag{12}$$

and

$$H_i=(e/\mu)\sum_i[\mathbf{p}_i\cdot\mathbf{A}_i-\mathbf{s}_i\cdot(\nabla\times\mathbf{A}_i)], \tag{13}$$

where $\mathbf{p}_i$ and $\mathbf{s}_i$ are the momentum and spin of ith electron, respectively. $\mathbf{A}_i$ is the vector potential at the position of ith electron. e and μ are the electron charge and mass, respectively. In H_i the contributions of nuclei are neglected.

In the same way as in (8) we set

$$\mathbf{A}_i=\sum_{jm}\int F(\omega jm)\mathcal{A}_{\omega jm}(\mathbf{r}_i)d\omega/c^{1/2}+\sum_{jm}\int F^*(\omega jm)\mathcal{A}_{\omega jm}^*(\mathbf{r}_i)d\omega/c^{1/2} \tag{14}$$

[8] A. I. Akhiezer and V. B. Berestetsky, *Quantum Electrodynamics* (izd. Fiz. Mat. Lit, Moscow, 1953) (English transl.: AEC-tr-2876).

[9] See for example, L. Pauling and E. B. Wilson, Jr., *Introduction to Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1935), Sec. 43.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in (13). If

$$\nu_1 = \nu - 1, \quad \nu_2 = \nu \tag{15}$$

in (11), we have

$$P(n_1\nu - 1 \leftrightarrow n_2\nu) = (\pi/\hbar^2)(\nu_{\omega_0 jm}{}^{(\mathrm{mg})}/c)\,|\,(n_1|\,(e/\mu)\sum_i [-\mathbf{p}_i\cdot\mathcal{A}_{\omega_0 jm} + \mathbf{s}_i\cdot(\nabla\times\mathcal{A}_{\omega_0 jm})]\,|\,n_2)^2, \tag{16}$$

where

$$\omega_0 = \omega_{n_1 n_2} \tag{17}$$

and $\nu_{\omega_0 jm}{}^{(\mathrm{mg})}$ is the number of the (magnetic jm) photons with frequency $\omega_0 \sim \omega_0 + d\omega$.

From (6) we see

$$\mathbf{p}_i\cdot\mathcal{A}_{\omega_0 jm} = (\hbar\omega/16\pi^3 c^2\epsilon_0)^{1/2} e^{-i\omega_0 t} g_j(\omega_0 r_i/c) \times r_i^{-2}(\mathbf{r}_i\times\mathbf{Y}_{jjm})\cdot\mathbf{l}_i, \tag{18}$$

where $\mathbf{l}_i$ is the orbital angular momentum of ith particle.

Using formulas

$$g_j r_i^{-2}(\mathbf{r}_i\times\mathbf{Y}_{jjm}) = -\nabla g_j Y_{jm}/[j(j+1)]^{1/2} \tag{19}$$

and

$$\nabla\times g_j\mathbf{Y}_{jjm} = \nabla g_j Y_{jm}[(j+1)/j]^{1/2}, \tag{20}$$

which can be shown from (4) we have

$$-\mathbf{p}_i\cdot\mathcal{A}_{\omega_0 jm} + \mathbf{s}_i\cdot(\nabla\times\mathcal{A}_{\omega_0 jm}) = (\hbar\omega_0/16\pi^3 c^2\epsilon_0)^{1/2} e^{-i\omega_0 t}\nabla g_n Y_{jm} \cdot[\{j(j+1)\}^{-1/2}\mathbf{l}_i + \{(j+1)/j\}^{1/2}\mathbf{s}_i]. \tag{21}$$

One can define the (magnetic jm) moment as

$$Q_{jm}{}^{(\mathrm{mg})} = (e/\mu)[(2j+1)/4\pi]^{1/2}(2j-1)!\,(i\omega_0/c)^{-j} \times\sum_i (\nabla g_j Y_{jm})_i\cdot[(j+1)^{-1}\mathbf{l}_i + \mathbf{s}_i]$$

$$\cong (e/\mu)[4\pi/(2j+1)]^{1/2} \times\sum_i (\nabla r_i{}^j Y_{jm})\cdot[(j+1)^{-1}\mathbf{l}_i + \mathbf{s}_i], \tag{22}$$

where the last expression is obtained when

$$\omega_0 r_i/c \ll 1, \tag{23}$$

so that

$$g_j(\omega_0 r/c) \cong 4\pi(i\omega_0 r/c)^j/(2j+1)!. \tag{24}$$

Note that

$$Q_{10}{}^{(\mathrm{mg})} = (e/\mu)\sum_i (l_{iz} + 2s_{iz}), \tag{25}$$

which is the familiar magnetic dipole moment.

From (21) and (22) we see

$$\sum_i [-\mathbf{p}_i\cdot\mathcal{A}_{\omega_0 jm} + \mathbf{s}_i\cdot(\nabla\times\mathcal{A}_{\omega_0 jm})](e/\mu) = [(j+1)/j(2j+1)]^{1/2}[(2j-1)!]^{-1} \times e^{-i\omega_0 t}(i\omega_0/c)^j Q_{jm}{}^{(\mathrm{mg})}/2\pi c. \tag{26}$$

Thus, from (16) the probability of the (magnetic jm) moment transition is

$$P_{jm}{}^{(\mathrm{mg})}(n_1 \rightarrow n_2) = (j+1)[j(2j+1)\{(2j-1)\}^2]^{-1}\nu_{\omega_0 jm}{}^{(\mathrm{mg})}\omega_0{}^{2j+1} \times(4\pi c^{2j+3}\hbar\epsilon_0)^{-1}\,|\,(n_1|Q_{jm}{}^{(\mathrm{mg})}|\,n_2)\,|^2. \tag{27}$$

ELECTRIC MULTIPOLE TRANSITION

The electric multipole transition can be discussed in the similar way. Since it is not the main purpose of the present paper, we will just show the result.

The only difference is to start with a photon wave function which is the eigenfunction of the total angular momentum and energy with eigenvalues j, m, and $\hbar\omega$, while the parity is $(-1)^j$, instead of $(-1)^{j+1}$.

The (electric jm) moment is defined as

$$Q_{jm}{}^{(\mathrm{el})} = \sum_i e_i[4\pi/(2j+1)]^{1/2} r_i{}^j Y_{jm}, \tag{28}$$

where we should include both electrons and nuclei in the summation, and e_i is the charge of ith particle. This definition gives

$$Q_{10}{}^{(\mathrm{el})} = \sum_i e_i z_i, \tag{29}$$

which is the familiar electric dipole moment.

The transition probability is

$$P_{jm}{}^{(\mathrm{el})}(n_1 \leftrightarrow n_2) = j(j+1)[(2j+1)\{(2j-1)!\}^2]^{-1}\nu_{\omega_0 jm}{}^{(\mathrm{el})}\omega_0{}^{2j+1} \times(4\pi c^{2j+1}\hbar\epsilon_0)^{-1}\,|\,(n_1|Q_{jm}{}^{(\mathrm{el})}|\,n_2)\,|^2. \tag{30}$$

Since the matrix element of $Q_{jm}{}^{(\mathrm{mg})}$ is about $(e/\mu)r_0{}^{j-1}\hbar$ as seen from (22), while that of $Q_{jm}{}^{(\mathrm{el})}$ is about $er_0{}^j$ as seen from (28), taking r_0 for the atomic dimension, we obtain the order of magnitude of $P_{jm}{}^{(\mathrm{mg})}$ by means of $P_1{}^{(\mathrm{el})}$ as

$$P_{jm}{}^{(\mathrm{mg})} = (\omega_0{}^{j-1}r_0{}^{j-2}\hbar/c^j\mu)^2(\nu_{\omega_0 jm}{}^{(\mathrm{mg})}/\nu_{\omega_0 1}{}^{(\mathrm{el})})P_1{}^{(\mathrm{el})}. \tag{31}$$

SPIN-ORBIT ELECTRIC OR MAGNETIC DIPOLE TRANSITION

The conventional theory of $\Delta S = \pm 1$ transition is to take into account the effect of the spin-orbit interaction

$$H_{\mathrm{so}} = -i(e\hbar/2\mu^2 c^2)\sum_i \mathbf{s}_i\cdot(\mathbf{E}\times\nabla)_i. \tag{32}$$

Since this interaction is similar to $Q_{2m_s}{}^{(\mathrm{mg})}$, which is given by (33), it can connect a singlet to a triplet. Thus,

$$|a00) + \sum_{c,M_s}(c1M_S|H_{\mathrm{so}}|a00)(E_a - E_c)^{-1}|c1M_S)$$
$$|b1M_S) + \sum_d (d00|H_{\mathrm{so}}|b1M_S)(E_b - E_d)^{-1}|d00) \tag{33}$$

are the improved wave functions of the initial and the final states. The electric dipole transition for $a \leftrightarrow b$ is

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

thus from (30)

$$P_{1mso}^{(el)}=(\nu_{\omega 01m}^{(el)}\omega_0^2/6\pi c^3\hbar^2 c_0)\times\Big|\sum_{c,M_s}(a00|H_{so}|c1M_S)(E_a-E_c)^{-1}\times(c|Q_{1m}^{(el)}|b)+\sum_{d,M_s}(d00|H_{so}|b1M_S)\times(E_b-E_d)^{-1}(a|Q_{1m}^{(el)}|d)\Big|^2. \quad (34)$$

Since the order of magnitude of the electric field E is (electric energy)$/r_0$ or $\mu\omega_0^2 r_0$, we see the order of magnitude of the matrix element of H_{so} is $\hbar\omega_0/\mu c^2$ as expected from the fact that the spin-orbit interaction appears as a relativistic correction. Thus, the transition probability due to this mechanism is

$$P_{1\,so}^{(el)}\cong(\hbar\omega_0/\mu c^2)^2 P_1^{(el)}. \quad (35)$$

From (31) we see that the spontaneous emission probability through the magnetic quadrupole transition, which is $P_2^{(mg)}$ with $\nu=1$, gives exactly the same expression as (35).

In some cases the electric dipole transition is still forbidden and only the magnetic dipole transition can appear with the spin-orbit coupling. The transition probability for such case has the order of magnitude given by

$$P_{1\,so}^{(mg)}\cong(\hbar\omega_0/\mu c^2)^2 P_1^{(mg)}. \quad (36)$$

From (31) we see

$$P_3^{(mg)}=(\omega_0^2 r_0^2/c^2)^2 P_1^{(mg)}=(\omega_0\mu\omega_0 r_0^2/\mu c^2)^2 P_1^{(mg)}\cong(\hbar\omega_0/\mu c^2)^2 P_1^{(mg)}, \quad (37)$$

which shows that the magnetic octupole transition is as important as the spin-orbit magnetic dipole transition.

Somehow the magnetic multipole transition has never been discussed, but it should be as important as the spin-orbit electric or magnetic dipole transition for $\Delta S=\pm 1$ transitions.

MAGNETIC QUADRUPOLE MOMENT

From (22) with $j=2$ we obtain the five components of the magnetic quadrupole moment. They are

$$Q_{2m}^{(mg)}=Q_{2ml}^{(mg)}+Q_{2ms}^{(mg)}, \quad (38)$$

$$Q_{2ms}^{(mg)}=(e/\mu)\sum_i[q_{iz}^{2m}s_{iz}+\tfrac{1}{2}q_{i+}^{2m}(s_{ix}+is_{iy})+\tfrac{1}{2}q_{i-}^{2m}(s_{ix}-is_{iy})]; \quad (39)$$

while $Q_{2ml}^{(mg)}$ is obtained by replacing s by l/3 in (33). Coefficients q's are tabulated in Table I.

Let us assume a two-electron system. Such system can be either a triplet or a singlet, if we neglect the spin-orbit interaction. The wave function of each state is a product of the orbital part and the spin part which can be designated as $|\xi SM_S)$. It is easy to see that

$$(a00|q_{1z}^{2m}s_{1z}+q_{2z}^{2m}s_{2z}|b10)=(\hbar/2)(a|q_{1z}^{2m}-q_{2z}^{2m}|b),$$
$$(a00|q_{1+}^{2m}s_{1+}+q_{2+}^{2m}s_{2+}|b1-1)=(\hbar/2)(a|q_{1+}^{2m}-q_{2+}^{2m}|b), \quad (40)$$
$$(a00|q_{1-}^{2m}s_{1-}+q_{2-}^{2m}s_{2-}|b11)=(\hbar/2)(a|q_{1-}^{2m}-q_{2-}^{2m}|b),$$

are the only finite matrix elements. We see that these matrix elements are not generally zero because the orbital part of the wave function $|a)$ and $|b)$ are symmetric and antisymmetric, respectively, under the exchange of two electrons. Since, as shown by Mizushima and Koide,[6] the corresponding matrix elements of all the electric multipole moments and the magnetic dipole moment are zero, the magnetic quadrupole moment is the lowest moment which has finite matrix elements between states with different total spin.

If we assume that

$$|a)=N[u(1)v(2)+v(1)u(2)],$$
$$|b)=N'[u'(1)v'(2)-v'(1)u'(2)], \quad (41)$$

where

$$N^{-2}=2\left[1+\left\{\int u(1)v(1)d\mathbf{r}_1\right\}^2\right],$$
$$N'^{-2}=2\left[1-\left\{\int u'(1)v'(1)d\mathbf{r}_1\right\}^2\right], \quad (42)$$

we obtain

$$(a|q_1-q_2|b)=2NN'[(v|v')(u|q|u')-(u|u')(v|q|v')+(u|v')(v|q|u')-(v|u')(u|q|v')]. \quad (43)$$

MAGNETIC OCTUPOLE MOMENT

The magnetic octupole moment $Q_{3m}^{(mg)}$ can also be divided into two parts:

$$Q_{3m}^{(mg)}=Q_{3ml}^{(mg)}+Q_{3ms}^{(mg)}, \quad (44)$$

$$Q_{3ms}^{(mg)}=(e/\mu)\sum_i[q_{iz}^{3m}s_{iz}+\tfrac{1}{2}q_{i+}^{3m}(s_{ix}+is_{iy})+\tfrac{1}{2}q_{i-}^{3m}(s_{ix}-is_{iy})], \quad (45)$$

while $Q_{3ml}^{(mg)}$ is obtained by replacing s by l/4 in (45).

Coefficients q's are tabulated in Table II.

For the matrix elements we see that formulas (40) and (43) holds for q^{3m}.

TABLE I. Coefficients q^{2m}.

m	q_z^{2m}	q_+^{2m}	q_-^{2m}
−2	0	0	$-6^{1/2}(x-iy)$
−1	$(3/2)^{1/2}(x-iy)$	0	$-6^{1/2}z$
0	$2z$	$x-iy$	$-x-iy$
1	$(3/2)^{1/2}(x+iy)$	$6^{1/2}z$	0
2	0	$6^{1/2}(x+iy)$	0

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Coefficients q^{3m}.

m	q_z^{3m}	q_+^{3m}	q_-^{3m}
−3	0	0	$-(45/4)^{1/2}(x-iy)^2$
−2	$(15/8)^{1/2}(x-iy)^2$	0	$-30^{1/2}z(x-iy)$
−1	$12^{1/2}z(x-iy)$	$(45/4)^{1/2}(x-iy)^2$	$-3^{1/2}(3z^2-r^2)$
0	$(9z^2-3r^2)/2$	$3z(x-iy)$	$-3z(x+iy)$
1	$12^{1/2}z(x+iy)$	$3^{1/2}(3z^2-r^2)$	$-(45/4)^{1/2}(x+iy)^2$
2	$(15/8)^{1/2}(x+iy)^2$	$30^{1/2}z(x+iy)$	0
3	0	$(45/4)^{1/2}(x+iy)^2$	0

HYDROGEN MOLECULE

The lowest triplet state of the hydrogen molecule is a dissociative state $^3\Sigma_u^+$. Since it is a dissociative state the lifetime of this state due to the spontaneous emission cannot be measured. The calculation in this section is just to show that the magnetic quadrupole transition can be much greater than the spin-orbit electric dipole transition.

The Heitler-London wave function[9] is well known to give a good approximation for the ground state $^1\Sigma_g^+$ and the lowest triplet state $^3\Sigma_u^+$. A large number of better wave functions are proposed for these states,[9] but for the present purpose of rough estimation the Heitler-London wave function will be good enough.

The Heitler London wave functions are obtained by identifying

$$\begin{aligned} u(i) &= u'(i) = \psi_{1s}(r_{ai}), \\ v(i) &= v'(i) = \psi_{1s}(r_{bi}), \end{aligned} \tag{46}$$

where r_{ai} and r_{bi} are distances between the ith electron and the protons a and b, respectively, and ψ_{1s} is the hydrogen wave function. From Table I it is easy to see that the only finite matrix elements of the magnetic quadrupole moment are

$$\begin{aligned} (^1\Sigma_g^+ 00 | Q_{20}^{(mg)} | ^3\Sigma_u^+ 10) &= 4(e\hbar/2\mu)N_+N_-r_{ab}, \\ (^1\Sigma_g^+ 00 | Q_{21}^{(mg)} | ^3\Sigma_u^+ 1-1) &= 24^{1/2}(e\hbar/2\mu)N_+N_-r_{ab}, \\ (^1\Sigma_g^+ 11 | Q_{2-1}^{(mg)} | ^3\Sigma_u^+ 11) &= 24^{1/2}(e\hbar/2\mu)N_+N_-r_{ab}, \end{aligned} \tag{47}$$

where r_{ab} is the internuclear distance. With the approximation of $N_+ = N_- = 2^{-1/2}$ we obtain

$$\begin{aligned} &P_{20}^{(mg)}(^1\Sigma_g^+ \leftrightarrow {}^3\Sigma_u^+ 0) \\ &\quad = (2/15)\nu\omega_0^5(4\pi c^7\hbar\epsilon_0)^{-1}(e\hbar/2\mu)^2 r_{ab}^2, \\ &P_{2\pm1}^{(mg)}(^1\Sigma_g^+ \leftrightarrow {}^3\Sigma_u^+ \pm 1) \\ &\quad = 5^{-1}\nu\omega_0^5(4\pi c^7\hbar\epsilon_0)^{-1}(e\hbar/2\mu)^2 r_{ab}^2, \end{aligned} \tag{48}$$

which shows that the magnetic quadrupole transition is more effective as the internuclear distance increases, as far as $\omega_0 r_{ab} < c$.

For the spin-orbit interaction one can divide the electric field $\mathbf{E}$ into two parts: The part due to the nucleus to which the electron under consideration is attached and the part due to the other atom. Assuming that the charge distribution is spherical around the other atom we have,

$$\mathbf{E} = \mathbf{E}_a + \mathbf{E}_b = (d\Phi_a/dr_{ai})\mathbf{r}_{ai}/r_{ai} + (d\Phi_b/dr_{bi})\mathbf{r}_{bi}/r_{bi}, \tag{49}$$

where ψ_a and Φ_b are potentials due to each part mentioned above. Now

$$\mathbf{r}_{ai} \times \nabla = (i/\hbar)\mathbf{l}_{ai}, \quad \mathbf{r}_{bi} \times \nabla = (i/\hbar)\mathbf{l}_{ai} + \mathbf{r}_{ba} \times \nabla, \tag{50}$$

where $\mathbf{l}_{ai}$ is the orbital angular momentum of the ith electron with respect to the nucleus a.

Using (46) we have

$$\mathbf{l}_{ai}u(i) = 0, \tag{51}$$

while

$$\begin{aligned} (1s | \mathbf{r}_{ba} \times \nabla | 2p_x) &= (16/2^{1/2}27)(r_{ba}/r_B)\mathbf{j}, \\ (1s | \mathbf{r}_{ba} \times \nabla | 2p_y) &= -(16/2^{1/2}27)(r_{ba}/r_B)\mathbf{i}, \end{aligned} \tag{52}$$

where $\mathbf{i}$ and $\mathbf{j}$ are unit vectors along the x and y axes, respectively, and r_B is the Bohr radius.

If the orbital part of the wave function of the $^1\Pi_u(1s2p\pi)$ state is

$$|^1\Pi_u) = [\psi_{1s}(1)\psi_{2p\pi}(2) + \psi_{1s}(2)\psi_{2p\pi}(1)]/2^{1/2}, \tag{53}$$

we find

$$\begin{aligned} &(^1\Pi_{u\pm} | H_{so} | ^3\Sigma_u^+ 1\pm1) \\ &\quad = (4/27)(e\hbar/2\mu^2)(d\Phi_b/dr_b)r_b^{-1}(r_{ba}/r_B)(\hbar/2) \\ &\quad = (Z/27\pi\epsilon_0)(e\hbar/2\mu c)^2(r_{ab}^2 r_B)^{-1}, \end{aligned} \tag{54}$$

where Z is effective charge of the second atom for the electron in the first atom. Z is less than one and goes to zero as r_{ab} increases. Since

$$(^1\Pi_{u\pm} | Q_{1\pm1}^{(el)} | ^1\Sigma_g^+) = (2^3/3^6)r_B e, \tag{55}$$

we see from (34) that

$$p_{1\pm1\,so}^{(el)} = (Z^2/10^7)\nu\omega_0^3 e^6\hbar^3(\mu^4c^7\epsilon_0^3)^{-1}(\Delta r_{ab}^2)^{-2}, \tag{56}$$

where

$$\Delta = E(^1\Pi_u) - E(^3\Sigma_u^+). \tag{57}$$

In obtaining this formula we assumed that the main contributions come from the mixing of the $^1\Pi_u(1s2p\pi)$ state to the $^3\Sigma_u^+$ state. Note that formula (56) shows that the transition probability in this case decreases as r_{ab} increases as $(Z/r_{ab}^2)^2$.

By comparing (56) to (48) we see

$$\begin{aligned} &P_2^{(mg)}/P_{1\,so}^{(el)} \\ &\quad = (10/Z^2)(\hbar\omega_0/Rch)^2(\Delta/Rch)^2(r_{ab}/r_B)^6, \end{aligned} \tag{58}$$

where R is the Rydberg constant. In our case

$$\begin{aligned} Z^2 &\sim 1/10, \\ \hbar\omega_0/Rch &\simeq 1/2, \\ \Delta/Rch &\sim 1/4, \\ r_{ab}/r_B &= 1.4, \end{aligned} \tag{59}$$

at the equilibrium distance of the ground state. Thus the above ratio is about 10 in favor of the magnetic quadrupole transition over the conventional spin-orbit electric dipole transition. When the internuclear distance r_{ab} increases the ratio will increase very rapidly.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

SELECTION RULES

The two comparative mechanisms, the magnetic quadrupole transition and the spin-orbit electric dipole transition have slightly different selection rules. For the case of $\Delta S=\pm 1$ transitions which we are interested in here the selection rules for the magnetic quadrupole transition are just like those of the electric dipole transition, or

$$\Delta L=0, \pm 1, \text{ but } L=0 \leftrightarrow 0; \text{ parity change}, \quad (60)$$

while those for the spin-orbit electric dipole transition are

$$\Delta L=0, \pm 1, \pm 2; \text{ parity change}, \quad (61)$$

in the case of atoms with L–S coupling. Thus, for example, $^1S \leftrightarrow {}^3S$ and $^2S \leftrightarrow {}^4D$ transitions can only be explained by the spin-orbit electric dipole transition.

Transitions with parity no change cannot be explained by either of them, and we have to go to the magnetic octupole transition or the spin-orbit magnetic dipole transition. Their selection rules for $\Delta S=\pm 1$ transitions are

$$\Delta L=0, \pm 1, \pm 2, \text{ but } L=0 \leftrightarrow 0; \text{ parity no change for magnetic octupole}, \quad (62)$$

$$\Delta L=0, \pm 1, \pm 2; \text{ parity no change for spin-orbit magnetic dipole}. \quad (63)$$

Selection rules for molecules can be found by finding those for the electric dipole and the electric quadrupole transitions. The selection rules for the magnetic quadrupole and the magnetic octupole transitions are the same as those respectively. For the benzene molecule, for example, the lowest triplet state $^3B_{1u}$ is known to be able to go to the ground state $^1A_{1g}$ by the spin-orbit electric dipole transition[5,6] but we see that the magnetic quadrupole transition is forbidden for this transition. In the case of the naphthalene and anthracene molecules, on the other hand, the lowest triplet state is assumed[10] to be $^3B_{2u}$, and the magnetic quadrupole transition is allowed from this state to the ground state $^1A_{1g}$. The lowest triplet state of the benzene molecule is observed[11] to have the lifetime of more than 300 sec, while in most aromatic molecules the triplet lifetime is accepted to be about 1 sec. This may be explained by the above conclusion, since our analysis on the hydrogen molecule showed that the magnetic quadrupole transition must be more important than the spin-orbit electric dipole transition in the case of large molecules.

[10] R. Pariser, J. Chem. Phys. 24, 250 (1956).
[11] D. P. Craig, J. M. Hollas, and G. W. King, J. Chem. Phys. 29, 974 (1958).

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Absolute Electron Excitation Cross Sections of Helium*

ROBERT M. ST. JOHN, FRANK L. MILLER,† AND CHUN C. LIN‡
Department of Physics, University of Oklahoma, Norman, Oklahoma
(Received 16 December 1963)

The absolute apparent electron excitation functions of helium have been measured for the $3\,^1S$, $4\,^1S$, $5\,^1S$, $6\,^1S$, $3\,^1P$, $4\,^1P$, $3\,^1D$, $4\,^1D$, $5\,^1D$, $6\,^1D$, $3\,^3S$, $4\,^3S$, $5\,^3S$, $3\,^3P$, $3\,^3D$, $4\,^3D$, $5\,^3D$, and $6\,^3D$ states at pressures sufficiently low so that the effects of radiation imprisonment and collisional excitation transfer can be neglected. Corrections due to polarization of the radiation and the cascading from the upper excited states have been applied to the experimental data and the true excitation functions obtained. The peak values and the shape of these excitation functions are compared with the results reported from other laboratories. Generally good agreement is found with the previous works where the collisional excitation transfer was properly reduced and allowance was made for the cascading effect. The experimental cross sections show satisfactory agreement with the theoretical values calculated by the Born approximation for the ^1S and ^1P states at high-electron energies, but are about four times larger than the theoretical values for ^1D states. In the case of the triplet series, the experimental cross sections exceed the calculated values by factors of ten to one hundred or more. It is concluded that the population of the triplet states is produced mainly by processes other than direct excitation.

I. INTRODUCTION

RECENT studies of electron excitation of helium atoms indicate that some of the atomic states may be populated to a large extent by collisional excitation transfer and cascading in addition to the usual electron-impact excitation process.[1,2] The excitation transfer and its concommitant effect on cascading may be reduced and even eliminated if measurements of the excitation are made with the helium gas at low pressure, i.e., about 1 μ or less. Furthermore the populations of the ^1P states are affected by imprisonment of resonance radiation at pressures above a few tenths of a micron and therefore

* Supported by the U. S. Air Force Office of Scientific Research, Grant AF–AFOSR–252–63.
† Present address: Fort Lewis A. & M. College, Durango, Colorado.
‡ Alfred P. Sloan Foundation Fellow.

[1] R. M. St. John and R. G. Fowler, Phys. Rev. 122, 1813 (1961).
[2] C. C. Lin and R. M. St. John, Phys. Rev. 128, 1749 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

measurements on them should be made below that onset pressure in order that the imprisonment mechanism not obscure the electron-impact excitation process. Cascading effects cannot be entirely removed by reduction of the gas pressure as excitation of the upper levels goes on at all pressures. Thus, in the determination of the electron impact excitation cross section, considerable care must be exercised in subtracting from the total observed (apparent) cross section the contributions from the various other excitation mechanisms. Indeed, the effects of collisional transfer[1-3] and imprisonment are partly responsible for the disagreement of the electron-excitation cross sections reported in the early literature as well as the discrepancy between them and theoretical cross sections.

With the recent improved experimental techniques it is possible to work at pressures low enough to suppress the effect of excitation transfer when observing all but the weakest transitions. In this paper we shall present the apparent excitation functions of 18 excited states of helium measured at low pressure. The cross sections for direct electron excitation are then obtained after the application of a series of corrections. The results of several experimental investigations will be compared. Additionally comparison of these cross sections with the theoretical values will be made and discussed.

II. EXPERIMENTAL METHOD

The apparatus for measuring the relative excitation functions has been described in a previous paper.[4] Absolute values of the excitation functions were obtained by calibration against a standard lamp as previously described.[4,5] The lamp was a tungsten ribbon filament pyrometer supplied and standardized by the General Electric Company. The lamp was operated at 3 temperatures in the range from 1400 to 1700°K and its emission calculated within a small wavelength interval for the wavelength of each transition observed. Emissivity tables for tungsten determined by Larrabee were used.[6] The sensitivity of the detection system was determined with the three-lamp temperatures. The maximum deviation of the sensitivity from the mean value at a given wavelength was generally less than 5%.

The solid angle viewed by the monochromator was determined by the use of a small variable size circular diaphragm which acted as the aperture stop. Transition probabilities used in this paper for evaluating the density of a given excited state by means of observation of the intensity of light emitted from it are those tabulated by Gabriel and Heddle.[7]

[3] C. C. Lin and R. G. Fowler, Ann. Phys. (N. Y.) 15, 461 (1961).
[4] R. M. St. John, C. C. Lin, R. L. Stanton, H. D. West, J. P. Sweeney, and E. A. Rinehart, Rev. Sci. Instr. 33, 1089 (1962).
[5] R. M. St. John, C. J. Bronco, and R. G. Fowler, J. Opt. Soc. Am. 50, 28 (1960).
[6] R. D. Larrabee, J. Opt. Soc. Am. 49, 619 (1959).
[7] A. H. Gabriel and D. W. O. Heddle, Proc. Roy. Soc. (London) A258, 124 (1960).

III. ANALYSIS

The equation relating population gain and loss rates per unit volume of the jth state of an atomic system is

$$\underset{\text{Electron impact gain}}{Q(j)[I_eN(g)/eS]}+\underset{\text{Cascade gain}}{\sum_i N(i)A(ij)}+\text{transfer gain}$$
$$=\underset{\text{Cascade loss}}{N(j)A(j)}+\text{transfer loss}, \quad (1)$$

where $Q(j)$ is the cross section for excitation by electron impact from the ground state to the jth state, I_e is the electron-beam current passing through the collision tube, S is the cross-sectional area of the electron beam, and e the electronic charge. The symbols $N(g)$ and $N(j)$ stand for the densities of the ground and jth states; $A(ij)$ is the probability of transition from the ith state to the jth state, while $A(j)$ is the total probability of transition from the jth state to lower states. In the above equation we have neglected the effect of imprisonment of radiation. This follows from the assumption that observations of ^1P functions, which are directly affected by imprisonment, are to be made at pressures low enough that imprisonment is negligible.

At low pressure the transfer excitation terms in Eq. (1) can be neglected. When this is done, one obtains

$$Q(j)[I_eN(g)/eS]+\sum_i N(i)A(ij)=N(j)A(j)\,. \quad (2)$$

It is customary to introduce the branching ratio B,

$$B(jk)=A(j)/A(jk)\,, \quad (3)$$

so that Eq. (2) can be rewritten as

$$Q(j)[I_eN(g)/eS]+\sum_i N(i)A(ij)=B(jk)N(j)A(jk)\,. \quad (4)$$

Here $N(j)\ A(jk)$ is the rate at which energy is emitted in connection with transition $j\rightarrow k$ and is determined experimentally from the photomultiplier current $I_P(\theta,j)$ measured at an angle θ relative to the electron beam. Since the radiation generally is not isotropic, a polarization correction factor $f_P(\theta j)$, is introduced so that $I_P(\theta j)f_P(\theta j)$ is proportional to the angular average of the light intensity and thus allows the determination of the entire light flux emitted from the collision chamber. Accordingly we have

$$N(j)A(jk)=C_1(j)I_P(\theta,j)f_P(\theta,j)\,, \quad (5)$$

where $C_1(j)$ is a proportionality constant determined from the calibration proceedure using the standard lamp. For the purpose of comparison with the previous data, we define $Q'(j)$ and $Q''(j)$ as

$$Q'(j)\frac{I_eN(g)}{eS}\equiv Q(j)\frac{I_eN(g)}{eS}+\sum_i N(i)A(ij)\,, \quad (6)$$

$$Q''(\theta,j)\equiv Q'(j)/f_P(\theta,j)$$
$$=eSC_1(j)I_P(\theta,j)B(j)/I_eN(g)\,. \quad (7)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

All the factors making up $Q''(\theta,j)$ are obtained by experimental observation except $B(j)$ which is determined from the theoretical transition probabilities.[7] In fact, $Q''(\theta,j)$ is the cross section which has been reported traditionally by observers of excitation functions. We shall call it the cross section uncorrected for polarization, cascade, and transfer effects; or simply the apparent cross section. Likewise, $Q'(j)$ can be regarded as the observed cross section which has been corrected for polarization but not for cascading and transfer effects, and is obtained from the experimental $Q''(j)$ cross section.

The polarization correction factor[8] is

$$f_P(\theta,j)=\tfrac{1}{3}\{[300-P(j)]/[100-P(j)\cos^2\theta]\}\,, \quad (8)$$

where $P(j)$ is the percentage polarization.

IV. EXPERIMENTAL RESULTS

The apparent excitation functions of 18 states as measured directly from the automatic processing apparatus are shown in Fig. 1. The graphs represent relative values. After the calibration procedure described earlier was followed, the absolute values of the apparent cross sections were obtained; the peak values are listed in Table I.

The apparent excitation function of the $3\,^1P$ level displayed in Fig. 1 was recorded at a pressure of 5.9 μ. The shape of this curve was quite independent of gas pressure. Data obtained at pressures as low as 0.1 μ were used to evaluate the peak apparent cross section of 350×10^{-20} cm^2 listed in Table I.

The excitation function obtained at 4025–6 Å represents the sum of two-line functions. These are 4026 Å ($5\,^3D\rightarrow 2\,^3P$) and 4025 Å ($7\,^1S\rightarrow 2\,^1P$). One can obtain an excellent approximation to the excitation of the $7\,^1S$ level by extrapolation from the 3, 4, 5, and 6 levels of that family; application of the appropriate branching ratio yields the function of the 4025-Å line, with a maximum value of 1.1×10^{-20} cm^2. This, when subtracted from the function representing the sum of the two excitation processes, yields the 4026-Å function which, in turn, yields the $5\,^3D$ apparent excitation function.

TABLE I. Maximum values of the apparent excitation cross sections (low-pressure values).

Level	Transition observed (Å)	Maximum cross section Q^{11} (10^{-20} cm^2)
$3\,^1S$	(7281)	49
$4\,^1S$	(5047)	24
$5\,^1S$	(4438)	9.2
$6\,^1S$	(4170)	4.8
$3\,^1P$	(5016)	350
$4\,^1P$	(3965)	159
$3\,^1D$	(6678)	42
$4\,^1D$	(4922)	17.6
$5\,^1D$	(4387)	9.0
$6\,^1D$	(4144)	4.7
$3\,^3S$	(7065)	107
$4\,^3S$	(4713)	35
$5\,^3S$	(4121)	12.3
$3\,^3P$	(3889)	97
$3\,^3D$	(5876)	31
$4\,^3D$	(4471)	12.0
$5\,^3D$	(4026)	6.2
$6\,^3D$	(3821)	3.9

V. PREVIOUS WORKS

The apparent electron excitation functions of helium have been measured by Lees,[9] Thieme,[10] Yakhontova,[11] McFarland and Soltysik,[12] and Heddle and Lucas.[13] Of these Lees, Thieme, and Yakhontova determined absolute cross sections. Additionally, Stewart and Gabathuler[14] have determined the peak values of the cross sections of several helium levels. Gabriel and Heddle[7] determined the cross sections for 108-eV electrons for a number of helium levels.

The observations by Thieme, Yakhontova, Heddle, and Lucas, Gabriel and Heddle, and Stewart and Gabathuler were at pressures low enough to greatly reduce transfer effects. Of these, Thieme and Stewart and Gabathuler, however, did not work at pressures low enough to eliminate imprisonment of resonance radiation, and thus their absolute determinations of the $3\,^1P$ cross sections are abnormally high. All observers worked at pressures low enough that light emission from ^1S, ^3S, and ^3P states was linear with pressure. All observers cited, including the authors of this paper, observed the excitation chamber in a direction normal to the electron beams. Thus all apparent cross sections determined at low pressures are subject to the same corrections required for polarization and cascade.

Table II shows the peak absolute values of the apparent cross sections of helium levels as measured by the authors of the paper, Yakhontova, Stewart and Gabathuler, Lees, and Thieme.[15] It also shows our values and those of Gabriel and Heddle at an electron energy of 108 eV.

[8] I. C. Percival and M. J. Seaton, Phil. Trans. Roy. Soc. (London) **A251**, 113 (1958).

[9] J. H. Lees, Proc. Roy. Soc. (London) **A137**, 173 (1932).

[10] O. Thieme, Z. Physik **78**, 412 (1932).

[11] V. E. Yakhontova, Vestn. Lening. Univ., Ser. Fiz. i Khim. **14**, 27 (1959). (Translation 951 by Atomic Energy Research Establishment, Harwell, Berkshire, England.)

[12] R. H. McFarland and E. A. Soltysik, Phys. Rev. **127**, 2090 (1962); **128**, 1758 (1962).

[13] D. W. O. Heddle and C. B. Lucas, Proc. Roy. Soc. (London) **A271**, 129 (1963).

[14] D. T. Stewart and E. Gabathuler, Proc. Phys. Soc. (London) **74**, 473 (1959).

[15] Some of these experimenters published values for cross sections for excitation of a given line transmitted from a given level. By use of the branching factor one can readily obtain the cross section for excitation to the upper level involved. We used the transition probabilities of Gabriel and Heddle for determining the branching factors and thus obtained level cross sections as displayed in Table II. Stewart and Gabathuler listed both line and level cross sections. Due to the fact that they used branching factors differing from those produced by the Gabriel and Heddle transition probabilities, their level cross sections show some slight variation from those we list for them.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

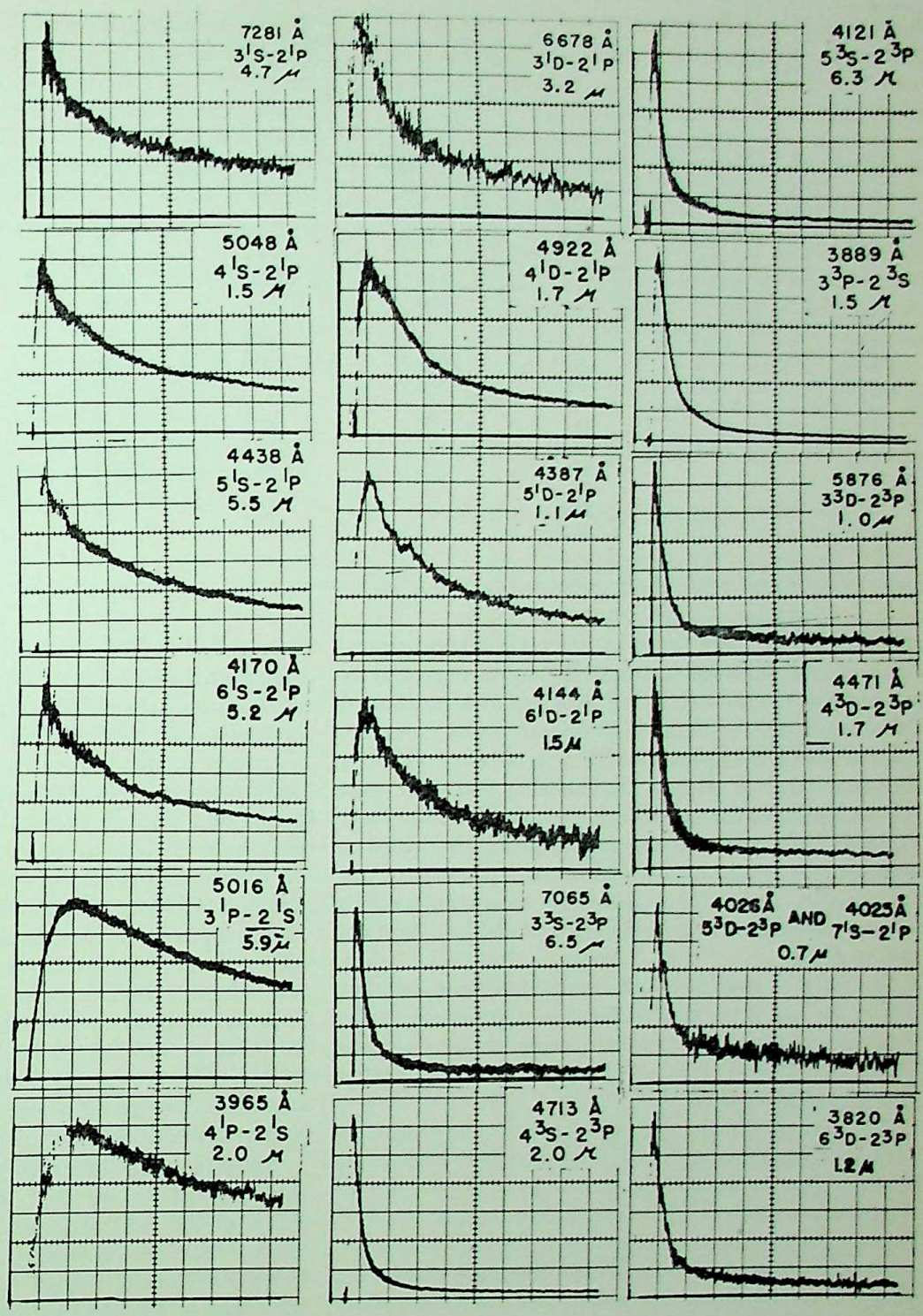


FIG. 1. Apparent excitation functions of 18 helium levels. Cross sections are relative; electron energy varies from 0 to 500 eV.

The measurements of Yakhontova are consistent with those reported herein. The ratio of Yakhontova's peak cross sections to ours averages 0.99 and each lies within the range of 0.83 to 1.16 with the exception of the $4\,^3S$ and $5\,^3D$ cross sections. In those cases the factors were 0.77 and 1.37. This rather good agreement throughout the levels including ^1D, and ^3D verifies that the pressures used in both investigations were below those which cause transfer effects.[16]

[16] Yakhontova determined the absolute value of the peak cross section of lines from the 3 and $4\,^1P$ levels by operating at a gas pressure high enough to cause complete imprisonment (which would yield a value near that of the level cross section) and then determined the line cross sections (5016 and 3965 Å) by dividing by the appropriate branching factor. Due to a spreading of the beam with increasing imprisonment which is concommitant with increasing pressure, the measured line intensity will not increase as much as the branching ratio when imprisonment changes from a nil to its full value. Thus, the line cross sections given by Yakhontova were not converted back to level cross sections and hence are not tabulated in Table II.

Stewart and Gabathuler do not quote directly the pressures used in their determinations of the apparent cross sections, but they were well aware of the pressure

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Comparison of absolute values of apparent cross sections of helium obtained by several experimenters. All cross sections are expressed in units of 10^{-20} cm^2.

Level	Cross section at peak: This paper	Yakhontova	Stewart and Gabathuler	Lees	Thieme	Cross section at 108 eV: This paper	Gabriel and Heddle
$3\,^1S$	49		27.5		28	31	36
$4\,^1S$	24	20	10.0	5.9	16.5	15	16.5
$5\,^1S$	9.2	7.7	6.1	2.8	9.5	6.1	7.0
$6\,^1S$	4.8		4130	4360	3660	3.0	4.0
$3\,^1P$	350		950	1270	1080	350	457
$4\,^1P$	159					158	210
$3\,^1D$	42		24	15.1	32	24	25
$4\,^1D$	17.6	17.8	12.2	9.3	18.6	12	12
$5\,^1D$	9.0	8.5	6.9		9.9	6.2	7.1
$6\,^1D$	4.7	5.0				3.3	3.0
$3\,^3S$	107		37	36	64	10.2	15
$4\,^3S$	35	25	17.2	8.6	40	3.3	4.4
$5\,^3S$	12.3	12.7	105	80	1890	1.17	1.44
$3\,^3P$	97	83	45		42	15.3	11
$3\,^3D$	31	36	18	15.2	23	4.3	25
$4\,^3D$	12.0	12.4	7.1			1.64	4.6
$5\,^3D$	6.2	8.5	4.3			0.89	3.0
$6\,^3D$	3.9					0.53	1.5

effects, as they gave intensity versus pressure curves. Some of that data was obtained in the 2-μ range, and we can assume the sensitivity of their equipment allowed them to obtain the maxima of the apparent cross sections at about that pressure. Their relative consistancy with our data for ^1D and ^3D indicates that they were able to minimize the transfer of excitation effects. Their values for $3\,^1P$ and $4\,^1P$ are very large and indicates an imprisonment effect as one expects as a pressure of 2 μ. Disregarding the ^1P data, the ratio of their cross sections to ours averaged a 1.26 and varied within the limits of 1.06 to 1.50.

The data of Lees is in serious disagreement with our results for the $3\,^1P$, due to his use of pressures in the 40-μ range. His peak apparent cross sections for the ^1S, ^1D, ^3S, ^3D, and ^3P levels compared to our values by a ratio averaging 0.86.

The data of Thieme compared to ours shows quite a variety of cross-section ratios; his extremely high value for excitation of the $3\,^1P$ level is at least in part due to imprisonment effects. His very large value for the $3\,^3P$ level defies explanation other than through erroneous calibration procedures. Lees' and Thieme's data were obtained by photographic techniques and hence more subject to errors than the photoelectric data of the other investigators cited. The magnitudes of Thieme's data will be given no further consideration.

Gabriel and Heddle measured absolute values of the cross section for 17 helium levels at pressures apparently low for all except the ^3D levels. The ratios of their apparent cross sections to those presented by the present authors averaged 1.12 for ^1S, ^1P, ^1D, ^3S, and ^3P levels, and lay within the range of 0.76 to 1.37. The ^3D ratio varied up to 3.4 indicating a transfer fillin for the cross sections of Gabriel and Heddle.

A comparison of the shapes of the excitation functions is presented in Table III. Each excitation function has been normalized to unity for an electron energy of 100 eV. Included are the data of Yakhontova, Lees, Thieme, McFarland and Solysik, Heddle, and Lucas, and that presented in this paper. It can be seen that there is quite good agreement between the shape data presented herein and that of Yakhontova. This is the result of both sets of data being obtained under low-pressure conditions. The shapes of the ^1D and ^3D curves of Lees are modified by the ^1P fillin due to his use of a relatively high gas pressure. The curves of Thieme are quite similar to those presented herein, a fact derived from the low-helium pressures used by him. McFarland and Soltysik obtained their functions with a gas pressure of 10 to 20 μ, and thus their curves show some transfer effects. The curves of Heddle and Lucas, while obtained at a pressure of about 1 μ, do not show very good agreement with the curves of the other experimenters.

Another factor influencing the shape of an experimentally determined excitation function is that of the shape of the electron beam. If it should undergo any change with electron energy, the light detection system being used could change in sensitivity due to (a) a variation in the light-gathering ability of the optical system with change in the position of the light source and (b) a lack of homogeneity of the response of in the cathode of the photomultiplier. Care was exercised by the authors of this paper to maintain a cylindrical electron beam of constant diameter at all electron energies.

VI. POLARIZATION AND CASCADE CORRECTIONS

Values of the percentage polarization, $P(j)$, have been measured by McFarland and Soltysik[12] for varying pressure and electron energy for several of the transitions reported in this paper, and are used for making

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the polarization corrections reported herein. Since the amount of polarization correction is generally rather small, the values of $P(j)$ given in Ref. 12 were used whenever available, and extrapolations of these values were used for the transitions which were not reported.

To obtain the true cross section of electron excitation, Q' must be corrected for the cascading effect. With the aid of Eqs. (5) and (7), Eq. (6) can be rewritten as

$$Q(j)=Q'(j)-\sum_i Q'(i)A(ij)/A(i)\,. \qquad (9)$$

The true excitation cross section of the jth state can then be determined from a knowledge of Q' of the jth state and of those states that cascade into it.

For our analysis we had available the values of $Q''(i)$ for those states included in Fig. 1 and Table I. These were corrected for polarization effects by the use of the polarization data of McFarland and Soltysik and values of $Q'(i)$ obtained. The shapes of high n values excitation functions corrected for polarization was taken to be the same as that of the low n functions having the same orbital quantum number L. The magnitudes of the high-level $Q'(i)$ functions were assumed to vary as $n^{-\alpha}$ where α is an adjustable parameter. Theoretical transition probabilities have been compiled for $n\leqslant 8$;[7] an extrapolation procedure again was used to find the values for higher states.

The excitation functions corrected for polarization and cascading effects of the levels of helium are shown in Fig. 2. The actual percentage of correction applied to an excitation function for each of the polarization and cascading effects varied with electron energy due to the differences between the energy dependence of the polarization factor, shape of the excitation function of the cascading levels, and the shape of the excitation function being corrected. The variation in the cascading correction was not so prominent for the triplet functions due to the similarity among triplet apparent excitation functions. The maximum percentage by which Q' varied from Q'' was 0 for ^1S levels, 4 to 7% for ^1P levels, 13 to 17% for ^1D levels, 0 for ^3S levels, 5% for the $3\,^3P$ level, and 4 to 5% for the ^3D levels. The higher levels showed less polarization than the low levels and thus received smaller corrections. This situation thus minimized the errors involved in the extrapolation procedure used in procuring the $Q'(j)$ curve from the $Q''(j)$ curve. The cascade contributions were then subtracted from the polarization corrected curves. The maximum cascade contribution to the $3\,^1S$ level was 15% and occurred at an electron energy of 300 eV. Cascade corrections of a similar percentage were obtained for higher ^1S levels. The $3\,^1P$ level sustained a 14% correction at 35 eV and a 4% correction at 100 eV. The $4\,^1P$ level received somewhat smaller percentage corrections and distinctly less absolute corrections. The ^1D functions suffered maximum corrections at 450 eV; they were in the range of 5 to 8%.

The cascade corrections sustained by triplet functions

TABLE III. Shape comparison of experimental apparent excitation functions of helium.

Level	Peak	This paper 60	This paper 100	This paper 200	This paper 400	Yakhontova Peak	Yakhontova 60	Yakhontova 100	Yakhontova 200	Lees Peak	Lees 60	Lees 100	Lees 200	Lees 400	Thieme Peak	Thieme 60	Thieme 100	Thieme 200	Thieme 400	McFarland and Soltysik Peak	McFarland and Soltysik 60	McFarland and Soltysik 100	McFarland and Soltysik 200	Heddle and Lucas Peak	Heddle and Lucas 60	Heddle and Lucas 100	Heddle and Lucas 200
$3\,^1S$	1.53	1.28	1.00	0.68	0.49																						
$4\,^1S$	1.55	1.26	1.00	0.66	0.42	1.33	1.08	1.00	0.79						2.02	1.40	1.00	0.66	0.41					1.6	1.2	1.00	0.77
$5\,^1S$	1.45	1.25	1.00	0.68	0.38	1.54	1.23	1.00	0.77	1.55	1.29	1.00	0.73	0.43	1.69	1.35	1.00	0.61	0.35								
$6\,^1S$	1.52	1.31	1.00	0.66	0.39					1.38	1.26	1.00	0.52		2.50	1.61	1.00	0.51	0.35								
$3\,^1P$	1.00	0.83	1.00	0.85	0.58	1.00	0.88	1.00	0.85	1.06	0.85	1.00	0.85	0.58	1.01	0.87	1.00	0.82	0.53	1.03	0.98	1.00	0.59	1.1	1.1	1.00	0.84
$4\,^1P$	1.00	0.76	1.00	0.87	0.60	1.00	0.86	1.00	0.84	1.16	0.94	1.00	0.85	0.60	1.01	0.85	1.00	0.79	0.42								
$3\,^1D$	1.73	1.50	1.00	0.51	0.30															1.08	1.06	1.00	0.63				
$4\,^1D$	1.34	1.32	1.00	0.47	0.28	1.49	1.37	1.00	0.60	1.28	1.13	1.00	0.70	0.47	1.70	1.40	1.00	0.57	0.55	1.40	1.35	1.00	0.44	1.75	1.48	1.00	0.56
$5\,^1D$	1.34	1.32	1.00	0.47	0.28	1.51	1.40	1.00	0.56	1.13	0.96	1.00	0.72	0.47	1.54	1.30	1.00	0.53	0.31	1.55	1.39	1.00	0.47				
$6\,^1D$	1.34	1.32	1.00	0.47	0.28	1.63	1.43	1.00	0.51						1.77	1.42	1.00	0.59	0.34								
$3\,^3S$	9.7	2.9	1.00	0.48	0.40																						
$4\,^3S$	9.7	2.9	1.00	0.48	0.40	10.6	2.6	1.00	0.67	20.8	2.7	1.00			14.6	2.62	1.00	0.18	0.14	14.9	3.6	1.00	0.29	14.8	2.0	1.00	0.77
$5\,^3S$	9.7	2.9	1.00	0.48	0.40	12.2	2.7	1.00	0.58	15.5	3.2	1.00			13.5	2.82	1.00	0.24	0.14								
$3\,^3P$	5.7	3.1	1.00	0.34	0.24	8.0	1.75	1.00	0.50	6.3	2.4	1.00	0.40	0.25	10.2	3.0	1.00	0.28	0.12	6.1	2.8	1.00	0.33	12.0	3.9	1.00	0.38
$3\,^3D$	6.9	2.7	1.00	0.65	0.48	4.0	1.46	1.00	0.63						4.42	1.95	1.00	0.52	0.38	2.38	1.33	1.00	0.76				
$4\,^3D$	6.9	2.7	1.00	0.65	0.48	4.9	1.85	1.00	0.69	1.10	1.09	1.00	0.76	0.5	4.30	1.78	1.00	0.5	0.28	2.46	1.50	1.00	0.55	12.0	2.75	1.00	0.47
$5\,^3D$	6.7	3.0	1.00	0.77	0.50	2.3	1.35	1.00	0.82																		
$6\,^3D$	6.9	2.7	1.00	0.65	0.48																						

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

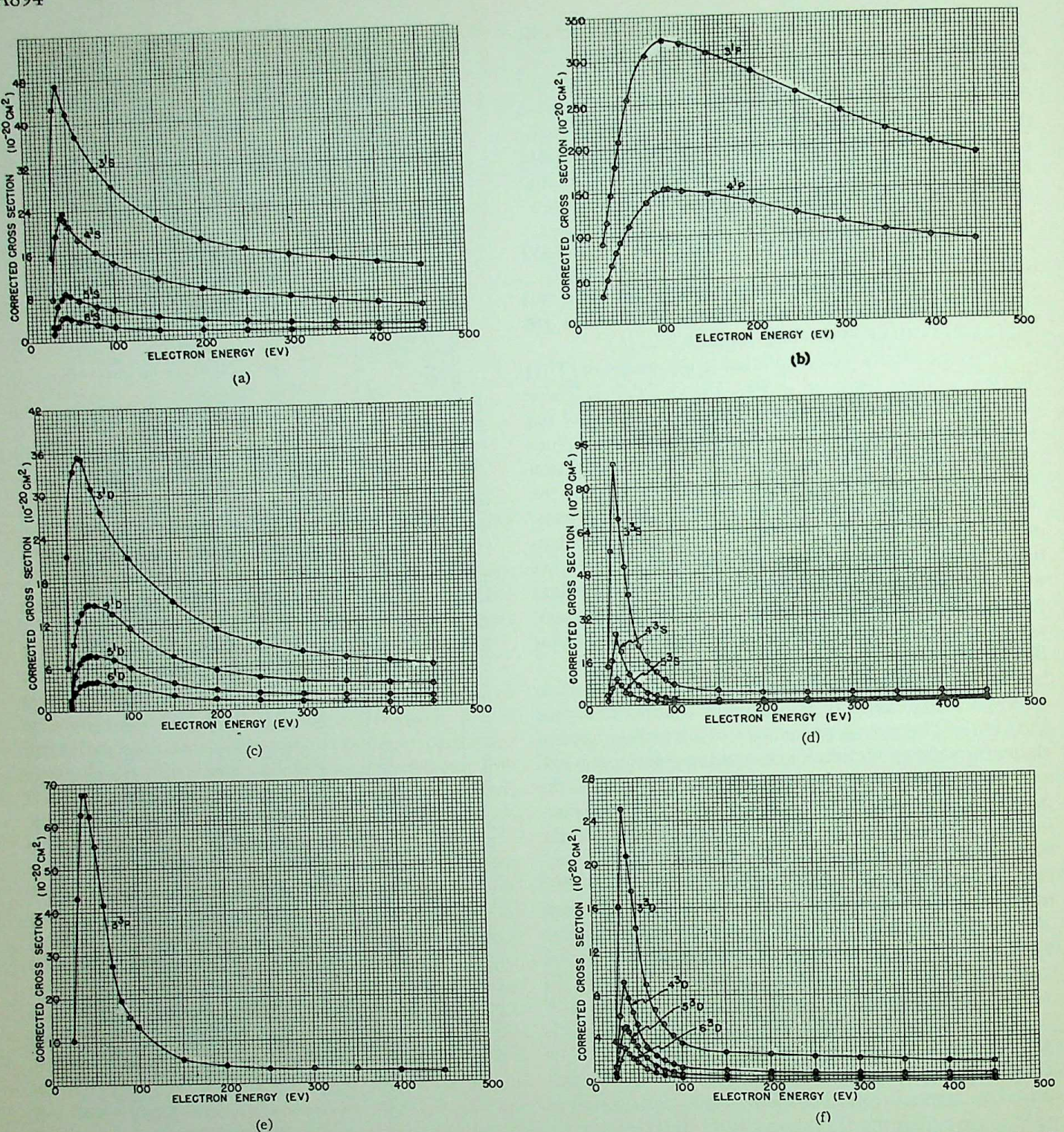


FIG. 2. Families of excitation function corrected for polarization and cascade effects. (a) Singlet S, (b) singlet P, (c) singlet D, (d) triplet S, (e) triplet P, (f) triplet D.

were nearly constant as a function of voltage. The $3\,^3S$, $4\,^3S$, and $5\,^3S$ functions of Q' were 20 to 30% due to cascade. The Q' curve of $3\,^3P$ was 30% due to cascade. The cascade correction for the ^3D functions varied to 10 to 19%.

Fortunately, the errors which might be introduced by the extrapolations processes necessarily are small since the cascading from high states is small. That part of the population of a state coming from levels involved in the extrapolation processes was 7% for the $3\,^1S$ state. The amounts falling to other low-level states by this process[17] were as follows: $3\,^1P$, 1%; $3\,^1D$, 4%; $3\,^3S$, 3%; $3\,^3P$, 4%; $3\,^3D$, 1%. There are additional uncertainties associated with the cascading from the F states. This point is discussed in Sec. VII.

[17] Measurements of the ^3P family were limited to $n=3$. The fall-off cross section with n was determined from the paper of Gabriel and Heddle which supplied cross section data for the 3, 4, and 5 states.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Frost and Phelps[18] made determinations of cascade contributions. They estimated that no correction was necessary for the ^1S and ^1D curves. They obtained a 10% correction for the $3\,^1P$ and $4\,^1P$ levels at 100 eV. The triplet curves, at the peak of the excitation function, were populated by cascade as follows: $4\,^3S$, 4%; $3\,^3P$, 49%; $3\,^3D$, $4\,^3D$, $5\,^3D$ all 10%.

Gabriel and Heddle[7] determined cascade corrections for electron energies of 108 eV. The $4\,^1S$ level population was 2% due to cascading. Higher ^1S level populations were regarded as entirely caused by electron impact. Singlet P populations were assumed free of cascade effects. The $3\,^1D$ level was found to have 4% of its population due to cascade, but higher ^1D unaffected by cascade. The triplets were populated by cascade as follows: $4\,^3S$, 23%; $5\,^3S$, 7%; $3\,^3P$, 34%; $3\,^3D$, 4%; $4\,^3D$, 17%; $5\,^3D$, 17%.

VII. COMPARISON WITH THEORY

The theory of excitation of helium atoms by electron impact has been treated rather extensively in the literature. Massey and Mohr[19] calculated the excitation cross sections for the $2\,^1S$, $3\,^1S$, $2\,^1P$ $3\,^1P$ $4\,^1P$, $5\,^1P$, $3\,^1D$, $4\,^1D$ $5\,^1D$ and $4\,^1F$ states by the Born approximation and for the $2\,^3S$, $3\,^3P$, $4\,^3P$, and $3\,^3D$ states using the Born-Oppenheimer approximation for incident electrons of various energies (60, 100, 200, and 400 eV). Curves for the theoretical excitation functions for the $3\,^1S$, $3\,^1P$, $4\,^1P$, $4\,^1D$, $5\,^1D$ and $2\,^3S$ states were also included in this work. Bates *et al.*[20] have given a critical analysis of the Born and Born-Oppenheimer approximation along with additional results of cross sections for some of the low-n states of He. As an improvement over the Born-Oppenheimer approximation, cross sections for the $1\,^1S \rightarrow 2\,^1S$ $1\,^1S \rightarrow 2\,^3S$, and $1\,^1S \rightarrow 2\,^3P$ transitions have been calculated by Massey and Moiseiwitsch[21] using the method of exchange distorted waves (EDW). Unpublished calculations by Fox on the excitation cross sections of several ^1S states were quoted in the paper of Gabriel and Heddle[7] and by Seaton.[22] In Table IV are listed observed and calculated (Massey and Mohr) excitation cross sections at electron energies of 60, 100, and 200 eV.

TABLE IV. Comparison of observed and calculated cross sections for electrons of various incident energies. Observed cross sections are corrected for polarization and cascading. All cross sections are expressed in units of 10^{-20} cm^2.

	60 eV		100 eV		200 eV	
Level	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
$3\,^1S$	38	26	28	15	18.5	11
$3\,^1P$	260	340	320	270	290	190
$4\,^1P$	112	140	152	110	136	76
$3\,^1D$	31		21	4.4	11.1	2.5
$4\,^1D$	14.7	3.7	11.6	2.5	5.6	1.3
$3\,^3S$	21.5	1.3[a]	7.6	0.60[a]	4.1	0.13[a]
$3\,^3P$	42	5.5	13.3	1.1	4.1	0.15
$3\,^3D$	9.0	0.13	3.5	0.03	2.6	0.004

[a] The $3\,^3S$ calculated values are extrapolated from $2\,^3S$ calculations.

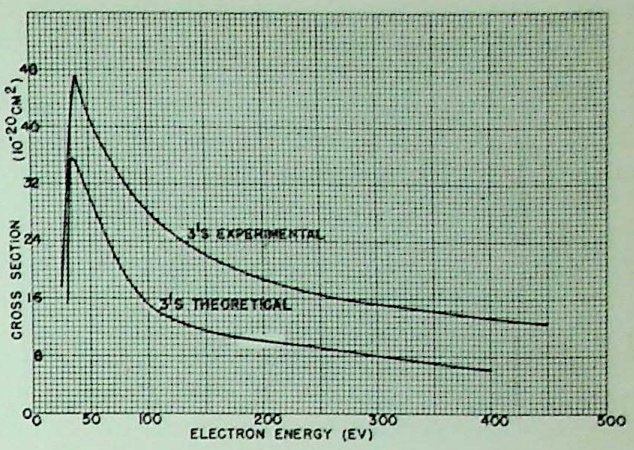


FIG. 3. Experimental and theoretical excitation functions of the $3\,^1S$ level. The experimental curve is corrected for polarization and cascade effects.

The theoretical and observed cross sections of the $3\,^1S$ level show good agreement. The calculated cross sections for the ^1S states are subject to a higher degree of uncertainty than those of the other states, because the wave functions of the $n\,^1S$ states were gotten by orthogonalization of the Slater-like orbitals. A small change of the effective nuclear charge, for instance, may result in a significant variation of the cross sections. Thus Massey and Mohr gave 15×10^{-20} cm^2 for the $3\,^1S$ cross section at 100 eV while the same cross section, according to the work of Fox,[23] is 46×10^{-20} cm^2. In the light of this, the experimental cross section may be considered to lie within the limits of the theoretical calculations. Figure 3 shows the comparison of the experimental and theoretical excitation functions of $3\,^1S$.

Figure 4 shows the comparison between the theoretical and experimental excitation functions of $3\,^1P$ and $4\,^1P$. The discrepancy at voltages below 100 eV is to be expected since the theoretical values were calculated by the use of the Born approximation. Analyses by Altshuler[24] and by Miller and Platzman[25] indicate that the use of approximate helium wave functions may lead to an error in the cross section as large as a factor of two or so. Thus the over-all agreement between the ob-

[18] L. S. Frost and A. V. Phelps, Westinghouse (Pittsburgh) Research Report 6-94439-6-R3, 1957 (unpublished).

[19] H. S. W. Massey and C. B. O. Mohr, Proc. Roy. Soc. (London) A140, 613 (1932).

[20] D. R. Bates, A. Fundaminsky, J. W. Leech, and H. S. W. Massey, Phil. Trans. Roy. Soc. (London) A243, 93 (1950).

[21] H. S. W. Massey and B. L. Moiseiwitsch, Proc. Roy. Soc. (London) A227, 38 (1954); A258, 147 (1960).

[22] M. J. Seaton, in *Atomic and Molecular Process*, edited by D. R. Bates (Academic Press Inc., New York, 1962).

[23] Fox's cross section of $3\,^1S$ were given for electron energy at 108 eV as 43×10^{-20} cm^2. The value at 100 eV is obtained by assuming the cross section to be inversely proportional to the energy of the colliding electron over this small range of energy.

[24] S. Altshuler, Phys. Rev. 87, 992 (1952); 89, 1093 (1953).

[25] W. F. Miller and R. L. Platzman, Proc. Phys. Soc. (London) A70, 229 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

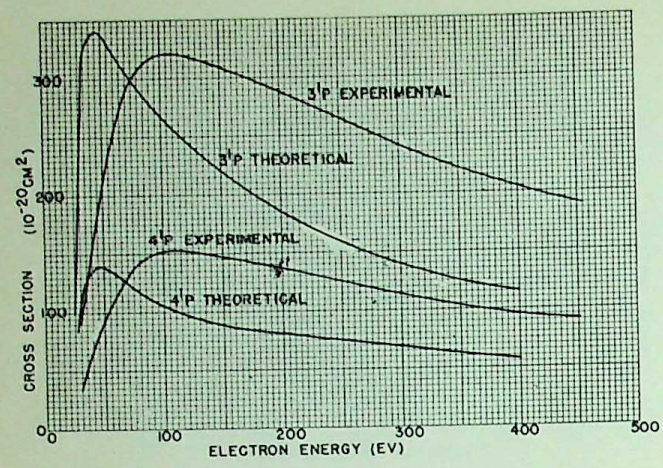


FIG. 4. Experimental and theoretical excitation functions of the $3\,^1P$ and $4\,^1P$ levels. The experimental curves are corrected for polarization and cascade effects.

served and calculated excitation cross sections for $3\,^1P$ and $4\,^1P$ should be considered as good as can be expected at this time.

The evaluation of the absolute cross sections Q from the experimental data for the ^1D states is somewhat complicated by the fact that the amount of cascading from the ^1F to the ^1D states are not known, since no excitation measurements are available for the ^1F states. We shall use the calculated value[19] of $Q(4\,^1F)$ along with an n^{-3} dependence for the higher ^1F states. The cascading from G to ^1F will be neglected, since the populations of the G states are expected to be very small on account of the rapid decrease of the direct excitation cross sections with increasing values of L. Under these assumptions we have calculated $N(n\,^1F)$ and found that the $^1F \rightarrow {}^1D$ cascading contributes less than 1% of the total population of the ^1D state. This result remains essentially the same when the singlet-triplet mixing of the F states are taken into consideration.[3] The absolute cross sections of 3, 4, 5, and $6\,^1D$ are then evaluated from the experimental data on Q'' (see Table I and Fig. 1). They are about four times larger than the theoretical values. At first, one might ascribe this discrepancy to the inaccuracy of our estimated populations of the F states. If we were to assume that $F \rightarrow {}^1D$ cascading is entirely responsible for the deviation of the observed Q from the theoretical values, it would be necessary to increase the concentration of the atoms in the F states. This, in turn, would require cross sections for electron excitation of the ^1F states to be 200 to 300 times as large as the theoretical values. The value of $Q(4\,^1F)$, for example, would have to be about 12×10^{-20} cm^2, a seemingly unreasonable value. Our experimental results for the excitation cross sections of the ^1D states therefore could not be brought into complete agreement with theory in a consistent manner.

Calculation by Fox as quoted by Seaton[22] gives $Q(3\,^1D)=7\times10^{-20}$ cm^2 at 108 eV as compared to 4×10^{-20} cm^2 calculated by Massey and Mohr[19] at 100 eV. The difference between these two values can be taken as an indication of the degree of variation of the theoretical cross section (Born approximation) which can be expected from the use of different approximate helium wave functions. Seaton has pointed out the calculations by B II approximation show that the coupling between $3\,^1D$ and $3\,^1P$ is not important for the cross section of the $3\,^1D$ state. It is very unlikely then that the imperfection of the theory could account for the disagreement between the observed and calculated cross sections for $3\,^1D$ and $4\,^1D$. Nor is it conceivable that errors in the experimental procedure could have caused the observed values to show this mismatch with the calculated cross sections. The measured apparent cross sections are proportional to the output current of the photomultiplier tube; thus, the relative values of the apparent excitation cross sections should be quite accurate. Of course the absolute experimental cross sections are subject to the errors of the light-intensity calibration as well as the cascade corrections. The amount of cascading populations in the ^1P and ^1D are quite small, i.e., less than 14% of the total populations. Errors in the calibration of the photomultiplier output should affect all the cross sections by a constant factor. This leads us to conclude that one cannot have simultaneous agreement between theory and experiment for the $3\,^1P$ $4\,^1P$, $3\,^1D$, and $4\,^1D$ states.

For the $3\,^3P$ and $3\,^3D$ states the calculated cross sections are consistently smaller than the observed values. In fact, the discrepancy is too large to be ascribed to the errors in the theory or experiment or combination of both. In the case of the ^3D states one is again faced with the problem of having to estimate the $F \rightarrow {}^3D$ cascading. Calculations show that the contributions of $F \rightarrow {}^3D$ cascading to the population of the $n\,^3D$ states are less than 1%, i.e., completely negligible.

Finally for the ^3S series, direct comparison between theory and experiment is more difficult since theoretical cross sections are available only for $2\,^3S$, while our experimental results cover $3\,^3S$, $4\,^3S$, and $5\,^3S$. In order to make any comparison it is necessary to extrapolate the experimental data with the assumptions that (i) the shape of excitation functions for all the $n\,^3S$ states is identical, and (ii) the peak values of these functions follow the relation $n^{-\alpha}$ or $(n^*)^{-3}$ as was assumed by Phelps and Frost[18] and by Gabriel and Heddle.[7] Here, n^* is the effective quantum number such that the energy of the state is equal to a negative constant divided by $(n^*)^2$. The extrapolation procedure yielded experimental values of $Q(2\,^3S)$ equal to 180, 69, and 39×10^{-20} at 60, 100, 200 eV, respectively, for the $n^{-\alpha}$ relations, where $\alpha=4.0$. Values of 140, 51, and 30×10^{-20} were obtained for the $(n^*)^{-3}$ relationship where $n^*=n-0.30$. These are to be compared with the theoretical values of 5, 2.4, and 0.5×10^{-20} cm^2 calculated by the Born-Oppenheimer approximation. Massey and Moiseiwitsch[21] have shown that when the distortion of the plane wave is taken into consideration, the calculated

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

cross sections become much smaller at low voltages. Thus it is seen that the experimental cross sections for the ^3S series are about 30 times the theoretical values. A comparison of the mismatch of theory and experiment of the ^3S family with that of the ^3P and ^3D can best be made at the $n=3$ level. Since no theoretical cross sections are available for the $3\,^3S$ state they were extrapolated from $n=2$ to $n=3$ by use of the $(n^*)^{-3}$ relationship. These figures are to be found in Table IV. They are used only to provide an estimate of the theoretical values and to draw conclusions of qualitative nature.

An examination of Table IV reveals that the fraction of the observed cross section accounted for by theory is much smaller for the $3\,^3D$ state than for the $3\,^3S$ and $3\,^3P$ states. However, a more important consideration would seem to be that of the magnitude of the unaccounted for cross sections, i.e., $Q(\text{obs})-Q(\text{theory})$. This value is of the same order of magnitude at a given electron energy for the three triplet states under comparison. Thus, the mechanism producing the excessive population in the triplet states does not preferentially populate any one state any more than the others.

VIII. CONCLUSIONS

The electron excitation cross sections for 18 excited states of helium have been determined by measuring the intensities of the radiation originated from these excited states at low pressure. Corrections have been made to allow for the cascading and polarization effects. The agreement between the observed and theoretical cross sections for the $3\,^1P$, $4\,^1P$, and $3\,^1S$ states is satisfactory. The experimental cross sections for $3\,^1D$ and $4\,^1D$ are about four times larger than the theoretical values. For the ^3S, ^3P and ^3D states the experimental results far exceeds (a factor of 10 or more) the calculated cross sections.

The discrepancy in the cross sections of the triplet series with the calculated values is far too large to be explained by experimental uncertainty. Nor can one reasonably expect the Born-Oppenheimer approximation to produce cross sections which are ten to a hundred times too small at electron energies above 60 or 100 eV. This is particularly true in view of the fact that the ^1P and ^1S cross sections calculated by the Born approximation do agree with the experimental values. Furthermore, calculations by Massey and Moseiwitsch[21] show that the excitation cross sections obtained from the Born-Oppenheimer approximation and from the EDW method approach each other at electron voltages above 50 eV. One does not expect the use of more refined method of calculation will change the theoretical cross sections (above 50 eV) significantly. In the case of $3\,^1D$ and $4\,^1D$ although the disagreement between theory and experiment is less severe, it is very unlikely that the cumulative errors of the theoretical calculation and the experimental work including the associated analysis of the excitation data could account for the deviation of a factor of four.

The explanation which seems most plausible to us is that the observed population of the triplet states (and possibly, to some extent, of the ^1D states) is produced mainly by processes other than direct excitation. Results of the measurements of the lifetime of the triplet states by Holzberlein and Fowler[26] also point toward the same conclusion. Of course, the mechanism or mechanisms with which the triplet states are populated must have the correct linear behavior with respect to pressure and electron beam current. Efforts are being made to perform additional experiments to provide further evidence of this "anomalous" behavior of the triplet states as well as to search for the mechanisms which are responsible for the observed populations of these states.

ACKNOWLEDGMENTS

The authors wish to thank R. J. Anderson, R. W. Kassik, and R. Sims for their assistance in the experimental work. Appreciation is also expressed to T. W. Nee for his part in processing the data.

[26] T. M. Holtzberlein and R. G. Fowler, Paper B-1, Sixteenth Annual Gaseous Electronics Conference, Pittsburgh, 1963 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

K-Shell Photoelectric Cross Sections from 200 keV to 2 MeV†

R. H. PRATT*
Institute of Theoretical Physics, Department of Physics, Stanford University, Stanford, California
AND
RICHARD D. LEVEE,‡ ROBERT L. PEXTON, AND WALTER ARON
Lawrence Radiation Laboratory, University of California, Livermore, California
(Received 13 September 1963)

A numerical program has been developed for the calculation of atomic photoelectric differential and total cross sections, including all polarization correlations. The program is designed to calculate relativistic Coulomb wave functions in a screened central potential; the outgoing continuum wave function is obtained in a partial-wave series. Results are presented here for K-shell differential and total cross sections in point Coulomb potentials (i.e., unscreened) ranging from charge $Z=13$ to $Z=92$, and covering the range of incident photon energies from 200 keV to 2 MeV. Enough data are presented to permit interpolation throughout these ranges. The total cross sections above 1 MeV are found to be significantly lower than previously accepted values. Further, the angular distributions from heavy elements deviate greatly from the commonly used Sauter distribution. These features are discussed and compared with existing experimental and theoretical work.

I. INTRODUCTION

ACCURATE predictions of the cross sections for the atomic photoelectric effect have in general been unavailable. Except for special limiting cases results must be obtained by numerical methods, even when the electron wave functions are assumed to be hydrogen-like and so available in analytic form. The complexity of these procedures has encouraged the use of extrapolation formulas based on various analytic approximations and indeed, at least for the total cross sections, moderate agreement with experiment has been obtained over a wide energy range.[1] The availability of accurate experimental total cross sections, the increasing need for accurate predictions of the differential cross sections, and the recent interest in the polarization properties of these reactions, have encouraged us to attempt their calculation by numerical means.[2]

We calculate in the central-field approximation, that is, we assume any atomic electron, whether bound or continuum, interacts only with a scalar spherically symmetric potential. Further, we neglect the effects of finite nuclear size, so that for sufficiently small distances this potential must simply reduce to the pure Coulomb Ze^2/r potential. For most applications these assumptions are appropriate.

Under these circumstances the theory of the atomic photoelectric effect is a simple application of first-order radiation theory. We require a fully relativistic treatment, since even at low energies, relativistic effects can be significant in a high-Z element. Then the differential cross section for the photoeffect is[1]

$$d\sigma/d\Omega=(2\pi)^{-2}p\epsilon|H|^2 \tag{1.1}$$

subject to energy conservation, where

$$H=-e(2\pi/k)^{1/2}\int d^3r\psi_{\rm fin}{}^*\boldsymbol{\alpha}\cdot\mathbf{e}e^{i\mathbf{k}\cdot\mathbf{r}}\psi_{\rm in}. \tag{1.2}$$

The absorbed radiation is described by its momentum $\mathbf{k}$ and polarization $\mathbf{e}$. $\psi_{\rm in}$ is a solution of the Dirac equation in a central potential corresponding to an initial bound state and $\psi_{\rm fin}$ is a solution corresponding to an outgoing electron of definite momentum $\mathbf{p}$ and energy ϵ ("plane wave" plus *ingoing* spherical waves).

Given the central potential, the entire problem is to solve the Dirac equation for the desired wave functions and then to integrate to obtain the matrix elements H. We have constructed a code for the IBM-7090 which does this.

In the present paper we present results for photoeffect from the K shell of an atom, obtained using wave functions in the pure point Coulomb field. The range of elements from aluminum ($Z=13$) to uranium ($Z=92$) was investigated for photon energies k from 200 keV to 2 MeV. Enough data were taken to permit interpolation throughout these ranges. For each choice of Z and k we obtained (1) the total cross section, (2) the differential cross section, and (3) all polarization correlations be-

† This study was supported in part by the U. S. Air Force Office of Scientific Research Grant AF-AFOSR-62-452, and in part by the U. S. Atomic Energy Commission.

* Address beginning August 1964: Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.

‡ Present address: Physics International, Berkeley, California.

[1] We will make some references to the previous work in this field in succeeding sections. The status of the subject up to 1954 is summarized for instance in W. Heitler, *Quantum Theory of Radiation* (Oxford University Press, New York, 1954), 3rd ed. We, in general, follow the notation of this book, but throughout we shall use the units $\hbar=c=m_e=1$. Thus, distances are measured in units of the electron compton wave length, etc. We will often use $a\equiv Ze^2$.

[2] While this work was in progress, an independent numerical calculation of K-shell photoelectric cross sections was reported by S. Hultberg, B. Nagel, and P. Olsson, Arkiv Fysik **20**, 555 (1961), hereafter referred to as HNO. This calculation, although using a method which cannot be applied to screened potentials and not obtaining results above 662 keV, is very valuable both for the information it provided on photoeffect at the lower energies and because it provides checking points for subsequent calculations, such as the present work.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tween incident photon and ejected electron. Results for differential and total cross sections are given here; the results obtained for the polarization correlations will be reported separately.[3] Most of the information is presented in tabular form. Modifications caused by screening,[4] extensions to lower and higher energies, and photoeffect from higher shells will be discussed in subsequent work.

The organization of the present paper is then as follows. In Sec. II we discuss the general properties of the photoeffect, as revealed by simple analytic approximations, and then develop the complete mathematical formalism for the process. Section III is devoted to a short discussion of the numerical methods for calculating wave functions, some features of which have not previously been published. The accuracy of the numerical calculations is discussed in Sec. IV. The remainder of the paper presents our results and compares them with experiment and previous theory. This is done for the total cross section in Sec. V and for the differential cross sections in Sec. VI.

II. GENERAL PROPERTIES AND FORMALISM

The main purpose of this section is to develop the formalism, based on Eqs. (1.1) and (1.2), needed for the calculation of the photoeffect. The wave function $\psi_{\rm fin}$ for the outgoing electron must be written as a sum over partial waves with appropriate asymptotic properties. Then for each partial wave the integrations over angles of Eq. (1.2) may be performed analytically, leaving a small number of radial integrals to be done numerically. The cross sections and the polarization correlations will be completely specified by these quantities.

It is useful, however, to first have a qualitative understanding of the general features of the process. For this purpose we replace Eq. (1.2) with the corresponding nonrelativistic expression

$$H=-e(2\pi/k)^{1/2}\int d^3r\psi_{\rm fin}{}^*\mathbf{p}\cdot\mathbf{e}e^{i\mathbf{k}\cdot\mathbf{r}}\psi_{\rm in}. \qquad (2.1)$$

We assume a pure Coulomb potential, so that for photoeffect from the K shell $\psi_{\rm in}=(a^3/\pi)^{1/2}e^{-ar}$, where $a\equiv Ze^2$. We also make the Born approximation and replace $\psi_{\rm fin}$ with the plane wave $e^{i\mathbf{p}\cdot\mathbf{r}}$. Then the integral (2.1) is easily done, and the differential cross section for photo effect from *both* K-shell electrons is given as

$$\frac{d\sigma}{d\Omega}=4e^2a^5(e_1e_1{}^*)\frac{p^3}{k^5\epsilon^3}\frac{\sin^2\theta}{(1-\beta\cos\theta)^4}, \qquad (2.2)$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{p}$, and e_1 is the component of the photon polarization vector in the scattering plane which $\mathbf{k}$ and $\mathbf{p}$ define.

This cross section vanishes both in the forward and backward directions. For low energies ($\beta\ll1$) it has a broad maximum centered at $\theta=90°$; as the energy increases the maximum shifts toward smaller angles and its width decreases. Equation (2.2) is independent of any circular polarization of the incident photons, but it is sensitive to linear polarization: If the beam is linearly polarized in the direction $\mathbf{e}$, $e_1=\cos\varphi$, where φ is the angle between the plane of $\mathbf{k}$ and $\mathbf{p}$ and the plane of $\mathbf{k}$ and $\mathbf{e}$; there is no emission perpendicular to the direction of polarization. Integrating over angles, the total cross section for unpolarized photons of low energy (but far above threshold) reduces to

$$\sigma=32\sqrt{2}\pi e^2a^5/3k^{7/2}. \qquad (2.3)$$

We note that this (1) varies as the *fifth* power of Z and (2) decreases rapidly with increasing energy.

These results must be modified both for low and for high energies. Near threshold the Born approximation expansion in a/β is invalid; when the exact nonrelativistic Coulomb wave functions are used it is found[1] that the total cross section (2.3) must be multiplied by

$$f=2\pi\left(\frac{I}{k}\right)^{1/2}\frac{e^{-4\eta\cot^{-1}\eta}}{1-e^{-2\pi\eta}},\quad \eta=[I/(k-I)]^{1/2}, \qquad (2.4)$$

where I is the ionization energy. This causes an appreciable reduction from the Born approximation prediction, which is reached only slowly as the energy increases, in the entire low-energy region; at the K-shell threshold $f=0.12$.[5]

For higher energies a relativistic treatment based on Eq. (1.2) is necessary. Sauter found that, to lowest order in Z, the differential cross section from linearly polarized photons is given by[6]

$$\frac{d\sigma}{d\Omega}=4e^2a^5\frac{p^3}{k^5\epsilon^4}\left[\frac{\sin^2\theta\cos^2\varphi}{(1-\beta\cos\theta)^4}-\frac{1-(1-\beta^2)^{1/2}}{2(1-\beta^2)}\frac{\sin^2\theta\cos^2\varphi}{(1-\beta\cos\theta)^3}+\frac{[1-(1-\beta^2)^{1/2}]^2}{4(1-\beta^2)^{3/2}}\frac{\sin^2\theta}{(1-\beta\cos\theta)^3}\right]. \qquad (2.5)$$

This again vanishes in the forward and backward directions; the maximum narrows and moves toward the forward direction as the energy increases. There is now some emission perpendicular to the direction of polariza-

[3] R. H. Pratt, R. D. Levee, R. L. Pexton, and W. Aron, Phys. Rev. 134, A916 (1964) (following paper).

[4] For K-shell photoeffect, it is believed that screening effects are quite small, typically (in heavy elements) of the order of 1–2%. For all other shells, screening effects are large and the use of hydrogen-like wave functions in a numerical calculation is not justified.

[5] However, in heavy elements relativistic effects are important at threshold. Part of this is simply due to the relativistic shift of the threshold energy. The relativistic K-shell photoeffect at threshold has been discussed by B. Nagel and P. Olsson, Arkiv Fysik 18, 29 (1960), in the approximation of an unscreened point Coulomb potential. However, as the authors note, screening effects are expected to be significant in the threshold region even for K-shell photoeffect.

[6] F. Sauter, Ann. Physik 11, 454 (1931).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tion. For very high energies the total cross section from unpolarized photons becomes

$$\sigma=4\pi e^2a^5/k, \tag{2.6}$$

and so still varies with the fifth power of Z but decreases less rapidly with increasing energy.

For high-Z elements the corrections to this lowest order theory can be very large. The high-energy behavior of the total cross section has been obtained exactly,[7] with numerical methods, and for $Z=82$ (lead) the prediction of Eq. (2.6) is too large by a factor of five. The expression

$$\sigma=[4\pi e^2a^5/k]e^{-\pi a}(1-4\pi a/15), \tag{2.7}$$

derived analytically, is in fair agreement with these results, and shows the nature of the suppression of the Z^5 dependence. It has also been shown that the cross section does not vanish in the forward direction when terms of relative order a^2 are considered, and such effects have been observed experimentally.[1]

In relativistic photoeffect there will also be interactions with the electron spin, and hence many more types of polarization correlations are possible. To describe these it is convenient to introduce the usual polarization parameters. We describe photon polarization with the quantities

$$\xi_1=e_1^*e_1-e_2^*e_2,\quad \xi_2=e_1e_2^*+e_2e_1^*,\quad \xi_3=i(e_1e_2^*-e_2e_1^*), \tag{2.8}$$

and we describe the polarization of the ejected electron by the direction ζ of its spin in its rest system. ζ_3 is taken along $\mathbf{p}$, ζ_1 in the scattering plane. Then if we sum or average over the initial polarization states of the bound electron, the differential cross section for photoeffect must be of the form

$$\frac{d\sigma}{d\Omega}=\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpol}}[\tfrac{1}{2}\sum_{i,j=0}^{3}\xi_i\zeta_jC_{ij}], \tag{2.9}$$

where $\xi_0=\zeta_0=C_{00}=1$, and $(d\sigma/d\Omega)_{\text{unpol}}$ is the differential cross section from unpolarized photons, summed over final electron spins. The correlation C_{10} connects linearly polarized photons and unpolarized electrons, and as already noted, occurs in the lowest order cross section. Also occurring in lowest order are C_{33}, connecting longitudinally polarized electrons and transversely polarized photons, and C_{31}. The correlations C_{02}, C_{12}, C_{21}, and C_{23} appear in relative order a. The remaining correlations are forbidden by invariance considerations.

We can now develop the formalism needed for a numerical calculation of these photoeffect cross sections. To begin with, we need wave functions for the electrons. A bound state will be specified by *capital* letters: total angular momentum J, orbital angular momentum L, $(J=L\pm\frac{1}{2})$, the azimuthal quantum number M, and total energy (including rest mass energy) E. Thus, for ψ_{in} in Eq. (1.2) we write

$$\psi_{\text{in}}=\begin{pmatrix}G_\kappa(r) & \Omega_{JLM}(\hat{r})\\ iF_\kappa(r) & \Omega_{JL'M}(\hat{r})\end{pmatrix}, \tag{2.10}$$

where L' and κ are defined by

$$L+L'=2J,\qquad \kappa=\mp(J+\tfrac{1}{2})\text{ as }J=L\pm\tfrac{1}{2},\quad(\text{or }J=L'\mp\tfrac{1}{2}), \tag{2.11}$$

and the spherical spinors Ω_{JLM} are defined in terms of spherical harmonics[8] by

$$\Omega_{JLM}=\begin{pmatrix}C_{JLM}^{+} & Y_{L,M-\frac{1}{2}}\\ C_{JLM}^{-} & Y_{L,M+\frac{1}{2}}\end{pmatrix}, \tag{2.12}$$

where

$$\begin{array}{lll} & J=L+\frac{1}{2} & J=L-\frac{1}{2}\\ C^{+}= & \left[\dfrac{J+M}{2L+1}\right]^{1/2} & \left[\dfrac{J-M+1}{2L+1}\right]^{1/2};\\ C^{-}= & \left[\dfrac{J-M}{2L+1}\right]^{1/2} & \left[\dfrac{J+M+1}{2L+1}\right]^{1/2}.\end{array} \tag{2.13}$$

The radial functions G_κ and F_κ are to be obtained as solutions of the coupled equations

$$\begin{aligned}&[E+1+\varphi]F_\kappa-[dG_\kappa/dr+(1+\kappa)(G_\kappa/r)]=0,\\ &[E-1+\varphi]G_\kappa+[dF_\kappa/dr+(1-\kappa)(F_\kappa/r)]=0.\end{aligned} \tag{2.14}$$

The potential φ must be specified (for the pure Coulomb potential $\varphi=a/r$) and the energy E is then determined from the eigenvalue problem. The wave function will be normalized by requiring

$$\int r^2dr[(F)^2+(G)^2]=1. \tag{2.15}$$

The wave function ψ_{fin} of the outgoing continuum electron is written as a sum over partial-wave solutions with appropriate asymptotic properties:

$$\psi_{\text{fin}}=\sum_{jlm}4\pi(\Omega_{jlm}{}^{+}(\hat{p})U_A)i^le^{-i\delta_\kappa}\begin{pmatrix}g_\kappa\Omega_{jlm}(\hat{r})\\ if_\kappa\Omega_{jl'm}(\hat{r})\end{pmatrix}. \tag{2.16}$$

Equations (2.10)–(2.14) are understood to remain valid with the substitution of the *small* letters j, l, l', m, κ, ϵ, etc., and of course $p^2+1=\epsilon^2$, where $\mathbf{p}$ is the momentum of the outgoing particle. The two-spinor U_A specifies the polarization properties of the electron in its rest frame. Equation (2.15), however, is not appropriate for continuum functions, and is replaced by the requirement that the functions are normalized such that at large

[7] R. H. Pratt, Phys. Rev. 117, 1017 (1960).

[8] We use the phase conventions of A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

distances

$$g_\kappa \underset{r\to\infty}{\longrightarrow} \left(\frac{\epsilon+1}{2\epsilon}\right)^{1/2}\frac{1}{pr}\sin(pr-\tfrac{1}{2}l\pi+\delta_\kappa),$$
$$f_\kappa \underset{r\to\infty}{\longrightarrow} \left(\frac{\epsilon-1}{2\epsilon}\right)^{1/2}\frac{1}{pr}\cos(pr-\tfrac{1}{2}l\pi+\delta_\kappa). \quad (2.17)$$

Equation (2.17) also defines the phase shifts δ_κ. [For the pure Coulomb potentials δ_κ in Eq. (2.17) but *not* Eq. (2.16) should be replaced by $\delta_\kappa+(a\epsilon/p)\ln 2pr$—this feature arises from the long-range character of the potential.] Note that the sum over j and l subject to $j=l\pm\frac{1}{2}$ in Eq. (2.16) is equivalent to a single sum over κ, where κ ranges over all positive and negative integers.

Substituting the wave functions (2.10) and (2.16) back into Eq. (1.2), H can be written as a summation of terms, each corresponding to a given κ and m in the series for the continuum wave function. It is convenient to remove some constants and write this in the form

$$H=-e(2\pi/k)^{1/2}\sum_\kappa \mathcal{H}_\kappa, \quad (2.18)$$

where each $\mathcal{H}_\kappa$ represents a sum over all m values consistent with that κ.

Now we choose a z axis along the photon direction $\mathbf{k}$ and start performing the angular integrations. In the φ integration nonvanishing contributions come only for $m=M\pm1$, and thus the summation over m implied in $\mathcal{H}_\kappa$ reduces to two terms. This is most conveniently written in terms of the circular polarization coefficients $e_\pm\equiv e_x\pm ie_y$, for then

$$\mathcal{H}_\kappa=4\pi e^{i\delta_\kappa}[(U_A{}^*\Omega_{jlM+1}(\hat{p}))e_-R_\kappa{}^+(M) + (U_A{}^*\Omega_{jlM-1}(\hat{p}))e_+R_\kappa{}^-(M)], \quad (2.19)$$

where the R's, a set of numbers resulting from θ and r integrations contain all the remaining information of the problem. (R's of $\pm M$ are easily related.) To proceed further for an arbitrary bound state one expands the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ of Eq. (1.2) in spherical harmonics and then integrates over products of three spherical harmonics.

For the purpose of this paper we restrict our attention to bound S states ($J=\frac{1}{2}$, $L=0$) and the results are quite simple. All the information of the problem is contained in two sets of numbers, $R_\kappa{}^\pm$, defined by

$$R_\kappa{}^+=\int r^2dr g_\kappa F_\kappa C_{jl+\frac{1}{2}}{}^+\left[\frac{l(l+1)}{2l+1}\right]^{1/2}\times[j_{l-1}(kr)+j_{l+1}(kr)],$$

$$R_\kappa{}^-=\int r^2dr\Big\{g_\kappa F_\kappa C_{jl-\frac{1}{2}}{}^- \times\left[-j_{l-1}(kr)\left(\frac{l^2}{(2l+1)}\right)^{1/2}+j_{l+1}(kr)\left(\frac{(l+1)^2}{2l+1}\right)^{1/2}\right] +\eta_\kappa f_\kappa G_\kappa C_{jl'-\frac{1}{2}}{}^-(2l'+1)^{1/2}j_{l'}(kr)\Big\}, \quad (2.20)$$

where $\eta_\kappa=\pm1$ according as $j=l\mp\frac{1}{2}$. [These R's correspond to $R^\pm(M=+\frac{1}{2})$ of Eq. (2.19).] Then for $M=+\frac{1}{2}$

$$\mathcal{H}_\kappa=4\pi e^{i\delta_\kappa}[(U_A{}^*\Omega_{jl+\frac{1}{2}}(\hat{p}))e_-R_\kappa{}^+ + (U_A{}^*\Omega_{jl-\frac{1}{2}}(\hat{p}))e_+R_\kappa{}^-], \quad (2.21)$$

and for $M=-\frac{1}{2}$

$$\mathcal{H}_\kappa=-4\pi\eta_\kappa e^{i\delta_\kappa}[(U_A{}^*\Omega_{jl-\frac{1}{2}}(\mathbf{p}))e_+R_\kappa{}^+ + (U_A{}^*\Omega_{jl+\frac{1}{2}}(\hat{p}))e_-R_\kappa{}^-]. \quad (2.22)$$

We shall *sum* over the two electrons of a bound S state. Then the differential cross section for a photon with polarization parameter ξ_i [as in Eq. (2.8)] to eject an electron with spin direction ζ (in its rest system, with ζ_3 chosen along the electron, and ζ_1 in the scattering plane) into the solid angle $d\Omega$ may be written

$$\frac{d\sigma}{d\Omega}=\tfrac{1}{2}A\sum_{i,j=0}^{3}\xi_i\zeta_jB_{ij}, \quad (2.23)$$

where

$$\xi_0=\zeta_0=1, \quad A=16\pi e^2p\epsilon/k. \quad (2.24)$$

The only nonvanishing B's are

$$\begin{aligned}
B_{00}&=[|J_-|^2+|K_-|^2+|J_+|^2+|K_+|^2],\\
B_{02}&=2\,\mathrm{Im}[J_-{}^*K_-+J_+{}^*K_+],\\
B_{10}&=2\,\mathrm{Re}[J_-{}^*J_++K_-{}^*K_+],\\
B_{12}&=2\,\mathrm{Im}[J_-{}^*K_++J_+{}^*K_-],\\
B_{21}&=-2\cos\theta\,\mathrm{Im}[J_-{}^*K_++K_-{}^*J_+] +2\sin\theta\,\mathrm{Im}[J_-{}^*J_++K_+{}^*K_-],\\
B_{23}&=-2\cos\theta\,\mathrm{Im}[J_-{}^*J_++K_+{}^*K_-] -2\sin\theta\,\mathrm{Im}[J_-{}^*K_++K_-{}^*J_+],\\
B_{31}&=2\cos\theta\,\mathrm{Re}[J_-{}^*K_--J_+{}^*K_+] -\sin\theta[|J_-|^2-|K_-|^2+|K_+|^2-|J_+|^2],\\
B_{33}&=\cos\theta[|J_-|^2-|K_-|^2+|K_+|^2-|J_+|^2] +2\sin\theta\,\mathrm{Re}[J_-{}^*K_--J_+{}^*K_+],
\end{aligned} \quad (2.25)$$

where

$$\begin{aligned}
J_-&=-(4\pi)^{-1/2}\sum_\kappa\eta_\kappa e^{i\delta_\kappa}\left[\frac{|\kappa-1|}{|\kappa||\kappa+1|}\right]^{1/2}R_\kappa{}^+P_l{}^1(\cos\theta),\\
K_-&=(4\pi)^{-1/2}\sum_\kappa e^{i\delta_\kappa}[|\kappa||\kappa^2-1|]^{-1/2}R_\kappa{}^+P_l{}^2(\cos\theta),\\
J_+&=(4\pi)^{-1/2}\sum_\kappa\eta_\kappa e^{i\delta_\kappa}[|\kappa|]^{-1/2}R_\kappa{}^-P_l{}^1(\cos\theta),\\
K_+&=(4\pi)^{-1/2}\sum_\kappa e^{i\delta_\kappa}[|\kappa|]^{1/2}R_\kappa{}^-P_l{}^0(\cos\theta).
\end{aligned} \quad (2.26)$$

In the sums of Eq. (2.26) terms for which the P_l's do not exist are to be omitted; θ is the angle between photon direction $\mathbf{k}$ and electron direction $\mathbf{p}$. [The last four B's of Eq. (2.25) take their form when a rotation is made so that ζ_3 refers to a spin along the *electron* direction.] To put our result in the form of Eq. (2.9) we finally define the polarization correlations by

$$C_{ij}\equiv B_{ij}/B_{00}. \quad (2.27)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The differential cross section from unpolarized photons, summed over electron spins is then

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpol}} = AB_{00} \tag{2.28}$$

and the total cross section

$$\sigma = 2\pi A \int_0^\pi B_{00} \sin\theta d\theta = A \sum_\kappa [(R_\kappa^+)^2 + (R_\kappa^-)^2]. \tag{2.29}$$

Thus, when the quantities $R_\kappa^\pm$ (and the phase shifts δ_κ) have been calculated, all properties of the process may be predicted. And the main work in calculating these quantities lies in the prior calculation of the electron wave functions.

III. METHOD OF SOLUTION OF THE WAVE EQUATIONS

A. The Bound-State Wave Equations

The bound-state wave equations

$$\begin{aligned}(E+1+\varphi)F_\kappa - [dG_\kappa/dr + (1+\kappa)(G_\kappa/r)] = 0,\\ (E-1+\varphi)G_\kappa + [dF_\kappa/dr + (1-\kappa)(F_\kappa/r)] = 0,\end{aligned} \tag{3.1}$$

constitute an eigenvalue problem for the energy eigenvalue E and the bound-state wave functions F_κ, G_κ. Since F_κ and G_κ behave like $r^{\gamma-1}$ near $r=0$ for the Coulomb potential[9] and are therefore singular for $\gamma<1$, it is convenient to make the substitution

$$G_\kappa = \bar{G}_\kappa r^{\gamma-1}, \quad F_\kappa = \bar{F}_\kappa r^{\gamma-1}.$$

Equations (3.1) then become

$$\begin{aligned}(E+1+\varphi)\bar{F}_\kappa - \left[\frac{d\bar{G}_\kappa}{dr} + \frac{(\gamma+\kappa)}{r}\bar{G}_\kappa\right] = 0,\\ (E-1+\varphi)\bar{G}_\kappa + \left[\frac{d\bar{F}_\kappa}{dr} + \frac{(\gamma-\kappa)}{r}\bar{F}_\kappa\right] = 0,\end{aligned} \tag{3.2}$$

where $\bar{F}_\kappa$ and $\bar{G}_\kappa$ are finite at $r=0$.

Although the following analysis for the solution of Eqs. (3.2) applies to a potential φ of any form, we shall in the present paper specify the Coulomb potential $\varphi = a/r$.

The boundary conditions associated with (3.2) with $\varphi = a/r$ are determined at the origin by multiplying through by r and letting $r=0$. We have then

$$\begin{aligned}a\bar{F}_\kappa(0) - (\gamma+\kappa)\bar{G}_\kappa(0) = 0,\\ a\bar{G}_\kappa(0) + (\gamma-\kappa)\bar{F}_\kappa(0) = 0.\end{aligned} \tag{3.3}$$

Since the equations are linear and homogeneous, we may choose $\bar{G}_\kappa(0)$ arbitrarily equal to one. $\bar{F}_\kappa(0)$ is then determined by either of the Eqs. (3.3), and γ is determined by the requirement that

$$\begin{vmatrix} a & -(\gamma+\kappa) \\ \gamma-\kappa & a \end{vmatrix} = 0$$

or

$$\gamma = (\kappa^2 - a^2)^{1/2}, \tag{3.4}$$

where the positive sign is taken to obtain a physically acceptable wave function.

In order to determine E we need to satisfy the boundary condition at $r\to\infty$; that is we must approach the asymptotic solution of (3.2). This solution is determined by letting $r\to\infty$ in (3.2), leading to

$$\begin{aligned}(E+1)\bar{F}_\kappa - \bar{G}'_\kappa = 0,\\ (E-1)\bar{G}_\kappa + \bar{F}'_\kappa = 0,\end{aligned}$$

which have the solution

$$\begin{aligned}\bar{G}_\kappa = C_1 e^{-\lambda r},\\ \bar{F}_\kappa = C_2 e^{-\lambda r},\end{aligned} \tag{3.5}$$

where $\lambda = (1-E^2)^{1/2}$ with $E<1$. Increasing solutions are ruled out since the wave function must be bounded as $r\to\infty$. [In the point Coulomb potential, these forms must be multiplied by $r^{n-|\kappa|}$, where n is the principal quantum number; this has no effect for the K shell.]

In order to numerically integrate (3.2), we must replace them by suitable difference equations. In the difference equations we have chosen, we evaluate the wave functions $\bar{F}$ and $\bar{G}$ (for simplification we have dropped the subscript κ) at the discrete points labeled by integral values of j running from zero at $r=0$ to J at a point r_J. The coordinate r_J is that point where the asymptotic boundary conditions (3.5) are satisfied in appropriate difference form to within some tolerance T_1. The conditions (3.2) are applied at $j+\frac{1}{2}$ and the derivatives are therefore evaluated as the slope of the chords joining values at j and $j+1$. Quantities needed at $j+\frac{1}{2}$ are evaluated by averaging. The resulting finite difference equations are

$$\begin{aligned}(E+1+\varphi_{j+\frac{1}{2}})\frac{(\bar{F}_{j+1}+\bar{F}_j)}{2} - \left[\frac{\bar{G}_{j+1}-\bar{G}_j}{r_{j+1}-r_j} + \frac{(\gamma+\kappa)}{2r_{j+\frac{1}{2}}}(\bar{G}_{j+1}+\bar{G}_j)\right] = 0,\\ (E-1+\varphi_{j+\frac{1}{2}})\frac{(\bar{G}_{j+1}+\bar{G}_j)}{2} + \left[\frac{\bar{F}_{j+1}-\bar{F}_j}{r_{j+1}-r_j} + \frac{(\gamma-\kappa)}{2r_{j+\frac{1}{2}}}(\bar{F}_{j+1}+\bar{F}_j)\right] = 0,\end{aligned} \tag{3.6}$$

with $0\le j\le J-1$.

The basis of the method is the simultaneous solution

[9] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., Sec. 44, p. 336.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the complete set of difference equations and boundary conditions. This method has been described for a two-point boundary value problem by Henyey *et al.*[10] The suggestion that the method could also be used for eigenvalue problems was made to one of us by Henyey[11] many years ago and has in fact been used by Levee to solve many problems.

The Eqs. (3.6), although linear in the functions $\bar{F}$ and $\bar{G}$, are nonlinear in the unknowns because of the appearance of the eigenvalue E. They must therefore be solved by an iterative technique such as the Newton-Raphson method described in Henyey *et al.*[10]

Formally we may write (3.6) as

$$\begin{aligned} &Q_{j+\frac{1}{2}}{}^1(\bar{F}_{j+1},\bar{F}_j,\bar{G}_{j+1},\bar{G}_j,E)=0, \\ &Q_{j+\frac{1}{2}}{}^2(\bar{F}_{j+1},\bar{F}_j,\bar{G}_{j+1},\bar{G}_j,E)=0, \\ &\qquad 0\leq j\leq J-1. \end{aligned} \tag{3.7}$$

The Newton-Raphson method when applied to a system

$$f_k(x_1,x_2,\cdots,x_K)=0, \quad k=1, 2, \cdots, K \tag{3.8}$$

corrects an approximate set of values $x_1{}^p$, $x_2{}^p$, $\cdots$, $x_K{}^p$ by variations $\delta x_1{}^p$, $\delta x_2{}^p$, $\cdots$, $\delta x_K{}^p$ to give improved values $x_1{}^{p+1}$, $x_2{}^{p+1}$, $\cdots$, $x_K{}^{p+1}$ where p is the iteration number. The variations are determined from the equations

$$f_k+\sum_{l=1}^{K}\frac{\partial f_k}{\partial x_l}\delta x_l=0, \quad k=1, 2, \cdots, K \tag{3.9}$$

and the iteration is continued until the Eqs. (3.8) are satisfied to within some desired tolerance.

If we define the following two vectors

$$Q_{j+\frac{1}{2}}=\begin{pmatrix} Q_{j+\frac{1}{2}}{}^1 \\ Q_{j+\frac{1}{2}}{}^2 \end{pmatrix}, \quad q_j=\begin{pmatrix} \delta\bar{F}_j \\ \delta\bar{G}_j \end{pmatrix},$$

we may write, in accordance with (3.9), the variational equations derived from (3.6) as

$$A_{j+\frac{1}{2}}q_{j+1}+B_{j+\frac{1}{2}}q_j+C_j\delta E+Q_{j+\frac{1}{2}}=0, \quad 0\leq j\leq J-1, \tag{3.10}$$

where

$$A_{j+\frac{1}{2}}=\begin{bmatrix} \frac{\partial Q_{j+\frac{1}{2}}{}^1}{\partial\bar{F}_{j+1}} & \frac{\partial Q_{j+\frac{1}{2}}{}^1}{\partial\bar{G}_{j+1}} \\ \frac{\partial Q_{j+\frac{1}{2}}{}^2}{\partial\bar{F}_{j+1}} & \frac{\partial Q_{j+\frac{1}{2}}{}^2}{\partial\bar{G}_{j+1}} \end{bmatrix}, \quad B_{j+\frac{1}{2}}=\begin{bmatrix} \frac{\partial Q_{j+\frac{1}{2}}{}^1}{\partial\bar{F}_j} & \frac{\partial Q_{j+\frac{1}{2}}{}^1}{\partial\bar{G}_j} \\ \frac{\partial Q_{j+\frac{1}{2}}{}^2}{\partial\bar{F}_j} & \frac{\partial Q_{j+\frac{1}{2}}{}^2}{\partial\bar{G}_j} \end{bmatrix},$$

$$C_j=\begin{bmatrix} \frac{\partial Q_{j+\frac{1}{2}}{}^1}{\partial E} \\ \frac{\partial Q_{j+\frac{1}{2}}{}^2}{\partial E} \end{bmatrix}.$$

Since the matrix of the coefficients of Eqs. (3.10) is of the form

$$\begin{array}{c} \\ j=0 \\ 1 \\ 2 \\ \vdots \\ J-1 \end{array}\begin{array}{c} \begin{array}{cccccccc} q_0 & q_1 & q_2 & q_3 & \cdots & q_{J-1} & q_J & \delta E \end{array} \\ \begin{bmatrix} x & x & & & & & & x \\ & x & x & & & & & x \\ & & x & x & & & & x \\ & & & \cdot & & & & \\ & & & & & x & x & x \end{bmatrix} \end{array},$$

we have J equations in $J+2$ unknowns. The two additional conditions which are necessary to solve the equations are given by the boundary condition at $j=0$ and the asymptotic condition at $j=J$.

At $j=0$, $\bar{F}_0$ is given by either of (3.3) when we have chosen $\bar{G}_0=1$ arbitrarily. Then $q_0=0$ and the first equation of (3.10) gives q_1 in terms of δE. Using this relation we eliminate q_1 from the second of (3.10) to give q_2 in terms of δE. The elimination process is continued to the final equation which gives q_J in terms of δE. These equations will be of the form

$$\begin{aligned} &q_j=e_j\delta E+f_j, \quad 0\leq j\leq J, \\ &q_0=0. \end{aligned} \tag{3.11}$$

Substituting the first of (3.11) into (3.10) we find recursion relations for the two-vectors e_j and f_j as

$$\begin{aligned} e_{j+1}&=-A_{j+\frac{1}{2}}{}^{-1}(B_{j+\frac{1}{2}}e_j+C_j), \\ f_{j+1}&=-A_{j+\frac{1}{2}}{}^{-1}(B_{j+\frac{1}{2}}f_j+Q_{j+\frac{1}{2}}). \end{aligned} \tag{3.12}$$

Since $q_0=0$, $e_0=f_0=0$ and we may determine all the e_j, f_j from (3.12) for $0\leq j\leq J$.

At $j=J$ we write the asymptotic solution (3.5) in difference form as

$$\bar{G}_J=\bar{G}_{J-1}e^{[-\lambda(r_J-r_{J-1})]}. \tag{3.13}$$

Taking the variation of (3.13) we have

$$\begin{aligned} \delta\bar{G}_J=e^{[-\lambda(r_J-r_{J-1})]}&[\delta\bar{G}_{J-1} \\ &-(r_J-r_{J-1})\bar{G}_{J-1}(d\lambda/dE)\delta E], \end{aligned} \tag{3.14}$$

where

$$d\lambda/dE=-E/(1-E^2)^{1/2} \tag{3.15}$$

from the asymptotic solution.

Now from (3.11) and our definition of q_j we have

$$\delta G_{J-1}=e_{J-1}{}^{(2)}\delta E+f_{J-1}{}^{(2)} \tag{3.16}$$

and

$$\delta G_J=e_J{}^{(2)}\delta E+f_J{}^{(2)},$$

where the superscript "2" refers to the lower element of the e, f two-vectors. Substituting (3.16) into (3.14) and

[10] L. G. Henyey, L. Wilets, K. H. Böhm, R. LeLevier, and R. D. Levee, Astrophys. J. 129, 628 (1959).
[11] L. G. Henyey (private communication to Levee).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

solving for E, we arrive at

$$\delta E=\frac{e^{-\lambda(r_J-r_{J-1})}f_{J-1}^{(2)}-f_J^{(2)}}{e_J^{(2)}+[\bar{G}_{J-1}(r_J-r_{J-1})(d\lambda/dE)-e_{J-1}^{(2)}]e^{-\lambda(r_J-r_{J-1})}}. \tag{3.17}$$

The iteration process may be summarized as follows: (1) From the pth approximation to $\bar{F}_j$, $\bar{G}_j$, and E, compute the elements of $A_{j+\frac{1}{2}}$, $B_{j+\frac{1}{2}}$, C_j and $Q_{j+\frac{1}{2}}$ for $0\leq j\leq J-1$. (2) Using (3.12) compute $e_1, e_2, \cdots, e_J$; $f_1, f_2, \cdots, f_J$ using $e_0=f_0=0$. (3) Compute δE by (3.17). (4) Compute the corrections $\delta\bar{F}_1, \delta\bar{F}_2, \cdots, \delta\bar{F}_J$, $\delta\bar{G}_1, \delta\bar{G}_2, \cdots, \delta\bar{G}_J$ from the first of (3.11). (5) Compute the $p+1$st approximations by

$$\left.\begin{aligned}E^{p+1}&=E^p+\delta E^p\\ \bar{F}_j{}^{p+1}&=\bar{F}_j{}^p+\delta\bar{F}_j{}^p\\ \bar{G}_j{}^{p+1}&=\bar{G}_j{}^p+\delta\bar{G}_j{}^p\end{aligned}\right\}0<j\leq J.$$

This process is repeated until all the unknowns satisfy the inequality

$$|\delta x_k{}^p/x_k{}^{p+1}|<T_2,\quad k=1, 2, \cdots, K$$

and $|\delta E/E|<T_1$ for $x_k>T_3$. For $x_k<T_3$ we do not make a test.

In addition to satisfying the above tolerance tests, we must also have chosen r_J such that the asymptotic solution is indeed valid. To check this we compute

$$\hat{\bar{F}}_{J-1}=\bar{F}_{J-2}e^{-\lambda(r_{J-1}-r_{J-2})}$$
$$\hat{\bar{G}}_{J-1}=\bar{G}_{J-2}e^{-\lambda(r_{J-1}-r_{J-2})}$$

and require that

$$|(\hat{\bar{F}}_{J-1}-\bar{F}_{J-1})/\bar{F}_{J-1}|<T_1$$

and that

$$|(\hat{\bar{G}}_{J-1}-\bar{G}_{J-1})/\bar{G}_{J-1}|<T_1.$$

If these inequalities are satisfied we are finished. If they are not, we advance J by 10 and resolve the equations. This procedure is repeated until all tests are satisfied.

In order to normalize our solutions we require that

$$N^2\int_0^\infty r^{\gamma+1}(\bar{F}^2+\bar{G}^2)dr=1=N^2\int_0^\infty H(r)dr.$$

This quadrature was performed by Bessell's formula

$$\int_0^\infty H(r)dr=\sum_{j=0}^{4}C_j(H_j+H_{J-j})+\sum_{j=5}^{J-5}H_j+\int_{r_J}^\infty H(r)dr$$

where the C_j take on the values

$$C_0{}^1=0.32986,\quad C_1{}^1=1.32083,\quad C_2{}^1=0.76667,$$
$$C_3{}^1=1.10139,\quad C_4{}^1=0.98125.$$

The final values of the wave functions for the bound state are given by

$$\left.\begin{aligned}F_j&=Nr^{\gamma-1}\bar{F}_j\\ G_j&=Nr^{\gamma-1}\bar{G}_j\end{aligned}\right\}0\leq j\leq J.$$

In the results reported here we have taken $T_1=10^{-4}$, $T_2=10^{-5}$, $T_3=10^{-5}$. The actual iteration error is much smaller since the tests involve comparison with values from the preceding iteration. In fact, since Newton's method doubles the number of significant figures when within the linear range, iteration errors are probably less than 10^{-8}.

B. The Continuum-State Wave Equations

The continuum-state wave equations

$$\begin{aligned}(\epsilon+1+\varphi)f_\kappa-[(dg_\kappa/dr)+(1+\kappa)(g_\kappa/r)]&=0\\ (\epsilon-1+\varphi)g_\kappa+[(df_\kappa/dr)+(1-\kappa)(f_\kappa/r)]&=0\end{aligned} \tag{3.18}$$

constitute an initial value problem since ϵ, the energy of the ejected electron, is determined by the sum of the incident photon energy and the energy of the bound state determined in the previous section.

As in the case of the bound state, it is convenient to make a change of variables here also.

$$f_\kappa=\bar{f}_\kappa r^{\gamma-1},\quad g_\kappa=\bar{g}_\kappa r^{\gamma-1}.$$

Equations (3.18) then become

$$\begin{aligned}(\epsilon+1+\varphi)\bar{f}_\kappa-[d\bar{g}_\kappa/dr+[(\gamma+\kappa)/r]\bar{g}_\kappa]&=0\\ (\epsilon-1+\varphi)\bar{g}_\kappa+[d\bar{f}_\kappa/dr+[(\gamma-\kappa)/r]\bar{f}_\kappa]&=0\end{aligned} \tag{3.19}$$

where $\bar{f}_\kappa$ and $\bar{g}_\kappa$ are finite at $r=0$.

In a manner similar to the bound state we determine at $r=0$ for a Coulomb potential a/r

$$a\bar{f}_\kappa(0)=(\gamma+\kappa)\bar{g}_\kappa(0) \tag{3.20a}$$

or

$$a\bar{g}_\kappa(0)=-(\gamma-\kappa)\bar{f}_\kappa(0). \tag{3.20b}$$

Thus $\gamma^2=\kappa^2-a^2$ and since Eqs. (3.20) are linear we may again choose $\bar{g}_\kappa(0)$ arbitrarily equal to 1. We then may determine $\bar{f}_\kappa(0)$ from either of (3.20). In practice we use (3.20a) for $\kappa>0$ and (3.20b) for $\kappa<0$ to insure the greatest accuracy.

In order to normalize the continuum solution properly, it is necessary to find the asymptotic solution. This is done by matching the solutions of (3.18) with φ not zero to general solutions of (3.18) with φ equal to zero,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and evaluating the asymptotic behavior from the zero-φ solutions beyond the range of the potential.

Let the matching of solutions be done at a point r_0. Let $p^2=\epsilon^2-1$. Then

$$\begin{aligned} f_\kappa(r_0)&=[(\epsilon-1)/pr_0]^{1/2} \\ &\quad\times[A_+J_{\kappa-\frac{1}{2}}(pr_0)+A_-J_{-(\kappa-\frac{1}{2})}(pr_0)] \\ g_\kappa(r_0)&=[(\epsilon+1)/pr_0]^{1/2} \\ &\quad\times[A_+J_{\kappa+\frac{1}{2}}(pr_0)-A_-J_{-(\kappa+\frac{1}{2})}(pr_0)]. \end{aligned} \tag{3.21}$$

These two equations serve to determine the constants A_+ and A_-. The asymptotic behavior of f_κ and g_κ is then determined by the asymptotic form of the Bessel functions.

$$\begin{aligned} f_\kappa(r)&\sim\left(\frac{2(\epsilon-1)}{\pi}\right)^{1/2}\frac{1}{pr}\left[A_+\cos\left(pr-\frac{\pi\kappa}{2}\right)\right. \\ &\quad\left.+A_-\cos\left(pr+\frac{\pi\kappa}{2}-\frac{\pi}{2}\right)\right], \\ g_\kappa(r)&\sim\left(\frac{2(\epsilon+1)}{\pi}\right)^{1/2}\frac{1}{pr}\left[A_+\cos\left(pr-\frac{\pi\kappa}{2}-\frac{\pi}{2}\right)\right. \\ &\quad\left.-A_-\cos\left(pr+\frac{\pi\kappa}{2}\right)\right]. \end{aligned} \tag{3.22}$$

The normalization factor A and the phase shift δ_κ are given by

$$\begin{aligned} A&=(A_+^2+A_-^2)^{1/2}, \\ \cos\delta_\kappa&=A_+/A, \\ -\sin\delta_\kappa&=(-1)^\kappa A_-/A. \end{aligned} \tag{3.23}$$

The asymptotic forms of f_κ and g_κ are then

$$\begin{aligned} f_\kappa(r)&\sim\left(\frac{2(\epsilon-1)}{\pi}\right)^{1/2}\frac{A}{pr}\cos\left(pr-\frac{\kappa\pi}{2}+\delta_\kappa\right), \\ g_\kappa(r)&\sim\left(\frac{2(\epsilon+1)}{\pi}\right)^{1/2}\frac{A}{pr}\sin\left(pr-\frac{\kappa\pi}{2}+\delta_\kappa\right). \end{aligned} \tag{3.24}$$

In using this with a true Coulomb potential, the phase shifts δ_κ will not become constant with increasing r, but rather will continue to increase, behaving like $(a\epsilon/p)\ln 2pr+$constant. This must be taken into account when checking whether the asymptotic form is attained. For potentials that die out more rapidly than the Coulomb potential, the δ_κ's should become constants in r.

The numerical solution of either (3.18) or (3.19) is straightforward except near $r=0$. Because of the singularity exhibited in (3.18) at $r=0$ we use (3.19), but even here we have a starting problem.

Let us write Eqs. (3.19) in the vector notation

$$\omega_\kappa'=A_\kappa\omega_\kappa, \tag{3.25}$$

where ω_κ has components $\bar{f}_\kappa$, $\bar{g}_\kappa$ and where

$$A_\kappa=\begin{bmatrix} -\dfrac{\gamma-\kappa}{r} & -\left(\epsilon-1+\dfrac{a}{r}\right) \\ \left(\epsilon+1+\dfrac{a}{r}\right) & -\dfrac{(\gamma+\kappa)}{r} \end{bmatrix}. \tag{3.26}$$

Consider now the finite difference representation of (3.25) based on the Euler approximation

$$(\omega_{j+1}-\omega_j)/h=A_j\omega_j, \quad j>0, \tag{3.27}$$

where $h=r_{j+1}-r_j=$a constant, and where we have for simplicity dropped the subscript κ. Equation (3.27) is restricted to $j>0$ since the elements of A_j are singular for $j=0$. In practice ω_1 would be determined from ω_0 by a series expansion. Equations (3.27) may be written

$$\omega_{j+1}=(I+hA_j)\omega_j.$$

The first-order error equation will then be

$$\delta\omega_{j+1}=(I+hA_j)\delta\omega_j, \tag{3.28}$$

where we have neglected errors in A_j. Equation (3.28) may be simplified to

$$\delta\omega_{j+1}=B_j\delta\omega_j, \tag{3.29}$$

with

$$B_j=\begin{bmatrix} 1-\dfrac{h}{jh}(\gamma-\kappa) & -h\left(\epsilon-1+\dfrac{a}{jh}\right) \\ h\left(\epsilon+1+\dfrac{a}{jh}\right) & 1-\dfrac{h(\gamma+\kappa)}{jh} \end{bmatrix} \tag{3.30}$$

and where we have replaced r by $r_0+jh=jh$.

If individual errors in ω_j are to remain bounded, we must require that the eigenvalue of the matrix B_j be less than or equal to 1 in absolute value. The eigenvalues of (3.30) are determined from

$$\begin{bmatrix} 1-\dfrac{(\gamma-\kappa)}{j}-\lambda & -h\left(\epsilon-1+\dfrac{a}{jh}\right) \\ h\left(\epsilon+1+\dfrac{a}{jh}\right) & 1-\dfrac{(\gamma+\kappa)}{j}-\lambda \end{bmatrix}=0.$$

The argument holds in the limit of $h\to 0$ which leads to

$$[1-(\gamma-\kappa)/j-\lambda][1-(\gamma+\kappa)/j-\lambda]+a^2/j^2=0. \tag{3.31}$$

Remembering that $\gamma^2=\kappa^2-a^2$ we solve (3.31) to give $\lambda=1,\ 1-(2\gamma/j)$. Since we require $|\lambda|\leq 1$ we have $-1\leq 1-(2\gamma/j)\leq 1$. Now γ and j are always greater than zero and therefore the right-hand inequality is always satisfied. The left-hand inequality will be satisfied if $j>\gamma$. The conclusion is that if we want errors to decrease in the solution of (3.19) by numerical integration, we must begin the integration at a value of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$j \geq \gamma$. The maximum value of γ in our calculations was 20 corresponding to $\kappa = \pm 20$ and it is therefore necessary to use the starting series described below for the first 20 steps. Although the above analysis is applied to the simple Euler method it leads to the same result where applied to the Runge-Kutta method. In fact, it was using the Runge-Kutta method for starting that led to the observations of the error growth and the analysis.[12] It has been pointed out to the authors by J. Scofield[13] that the modified Euler scheme, $\omega_{j+1} = [(I + \frac{1}{2}hA_{j+\frac{1}{2}})/(I - \frac{1}{2}hA_{j+\frac{1}{2}})]\omega_j$ has eigenvalues $\lambda = 1$, $[1 - \gamma/(j+\frac{1}{2})]/[1 + \gamma/(j+\frac{1}{2})]$ in the limit as $h \to 0$. Therefore errors in this scheme will be bounded since $\gamma > 0$.

The power-series solution to (3.19) around the point r_n is of the form

$$\begin{aligned} \bar{f}(r) &= \sum c_i (r - r_n)^i, \quad i = 0, 1, 2, \cdots, \\ \bar{g}(r) &= \sum b_i (r - r_n)^i, \quad i = 0, 1, 2, \cdots. \end{aligned} \tag{3.32}$$

Substituting into (3.19) we arrive at the recursion relations

$$\begin{aligned} r_n(i+2)b_{i+2} + (i+1)b_{i+1} + (\gamma+\kappa)b_{i+1} \\ - [(\epsilon+1)r_n + a]c_{i+1} - (\epsilon+1)c_i = 0, \\ r_n(i+2)c_{i+2} + (i+1)c_{i+1} + (\gamma-\kappa)c_{i+1} \\ + [(\epsilon-1)r_n + a]b_{i+1} + (\epsilon-1)b_i = 0, \end{aligned}$$

where $c_0 = \bar{f}(r_n)$, $b_0 = \bar{g}(r_n)$. These relations allow the determination of the c_i, b_i for $i > 0$ for successive expansions about $r_n = 0$, $r_n = h$, $r_n = 2h$, $\cdots$, $r_n = jh$. Since we are always expanding about the previous point, $r - r_n = h$.

The series solutions (3.32) and similar equations for the derivatives of $\bar{f}$ and $\bar{g}$ were computed out to $j = JJ$. The series at each j were terminated when the percentage error in f' and g' was less than one part in 10^8. Using the functions and derivatives thus determined, the numerical solution of (3.32) was then continued by the Runge-Kutta method. In the case of a general potential where the series solution is not easily arrived at, we plan using the modified Euler scheme for starting the integration.

At this point in j we switched back to using f and g instead of $\bar{f}$ and $\bar{g}$ so that our numbers would not become too small. In addition it was necessary to add a scale factor, SF, in both f and g in certain cases to keep numbers within machine limits. Using the values

$$\begin{aligned} f_{JJ} &= \bar{f}_{JJ}(SF \cdot r_{JJ})^{\gamma-1} \\ g_{JJ} &= \bar{g}_{JJ}(SF \cdot r_{JJ})^{\gamma-1} \\ f_{JJ'} &= \bar{f}_{JJ'}(SF \cdot r_{JJ})^{\gamma-1} + (\gamma-1)f_{JJ}/r_{JJ} \\ g_{JJ'} &= \bar{g}_{JJ'}(SF \cdot r_{JJ})^{\gamma-1} + (\gamma-1)g_{JJ}/r_{JJ}, \end{aligned}$$

the numerical integration of (3.32) was continued by the standard Runge-Kutta method.

[12] This difficulty had been observed and resolved previously by Levee in a problem in controlled thermonuclear reactions.

[13] J. Scofield (private communication).

Although our analysis in this section indicates that we should take $JJ = 20$, we actually used $JJ = 10$. This allowed small errors to appear in the normalization factors and wave functions for $\kappa > 10$. These errors were of the order of 4% in the normalization factor at $\kappa = 16$. However, the high values of κ contributed very little to the cross sections and the cross sections show no error due to using $JJ = 10$ instead of 20.

We require f and g to be normalized such that

$$\begin{aligned} f &\sim \left(\frac{\epsilon-1}{2\epsilon}\right)^{1/2} \frac{1}{pr} \cos\left(pr - \frac{l\pi}{2} + \delta\right) \\ g &\sim \left(\frac{\epsilon+1}{2\epsilon}\right)^{1/2} \frac{1}{pr} \sin\left(pr - \frac{l\pi}{2} + \delta\right). \end{aligned} \tag{3.33}$$

From (3.23) and (3.33) we then have the following relation for the normalization factor N_1: $N_1 = 2A(\epsilon/\pi)^{1/2}$ where A is determined from (3.21) and the first of (3.23).

The phase δ must now be determined. For a potential which falls off more rapidly than $1/r$ the phase will be given by (3.23) as $\delta = \arctan|A_-/A_+|$, but for the pure Coulomb potential $\delta = \arctan|A_-/A_+| - (a\epsilon/p)\ln 2pr$.

In the program reported here we have computed A at each step of the integration and required that

$$\left| 2\frac{A_{j+10} - A_j}{A_{j+10} + A_j} \right| < T_4.$$

It turned out that for the very small T_4 chosen, the integration was always carried out to the predetermined maximum at $r = 120$. At this point the normalizations and phases were determined.

IV. ACCURACY OF THE NUMERICAL CALCULATION

We have seen that we will have all the information needed to specify the properties of S-state photoeffect when we have calculated the quantities $R_\kappa^\pm$ given by Eq. (2.20). We have not discussed how many of these quantities we need to calculate, i.e., the convergence properties of the series in Eqs. (2.24) and (2.27). We follow usual practice, and will judge the convergence by the size of successive $R_\kappa^\pm$. Practical considerations force us to choose a limit for $|\kappa|$ in advance, since the methods employed for calculating functions and organizing the numerical program for machine purposes depend on this choice; in the present paper we consider $|\kappa| \leq 20$.

Then first of all the calculation of a given R_κ requires bound-state radial functions F_κ, G_κ, continuum radial functions f_κ, g_κ, and spherical Bessel functions $j_l(kr)$. Also, in calculating angular distributions we will need the associated Legendre functions $P_l^m(\cos\theta)$, for $m = 0$, 1, 2. The accuracy of the calculations of these *components* of the general program are discussed in separate subsections. We can then estimate the accuracy with which each R_κ has been computed, and the accuracy this implies for the cross sections. And we must also judge

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the error which has been made due to the restriction $|\kappa| \leq 20$.

Checks for the component programs will be discussed separately. A number of checks are available for the final results: We can compare with analytic results for low Z and we can compare with numerical results which have previously been obtained for photoelectric cross sections. These comparisons are made in the subsequent sections where we present our results. In all cases the proper agreement was obtained, so that we have full confidence in our numerical program. It should perhaps also be mentioned that with the present program some ten minutes of IBM-7090 time were required to obtain the differential and total cross sections and polarization correlations for a given choice of element and energy; it is expected that further improvements in the program will appreciably reduce this figure.

A. Component Programs

i. Bound-State Wave Functions

The accuracy of the bound-state calculation can be tested for the pure Coulomb potential by comparison with analytic results. It is particularly simple to check the eigenvalue and the normalization of a calculated wave function. The energy levels are obtained with considerable accuracy. However, the normalization coefficient, i.e., the constant which multiplies the known small r dependence, probably gives a better indication of the over-all accuracy of the wave function. For given $h \equiv r_{i+1} - r_i$, the error in mormalization increases with Z. With $h=0.05$, the smallest value used, the error is negligible for $Z=26$ or 50 but 0.4% for $Z=82$ and 0.6% for $Z=92$. To obtain 1% accuracy in the cross sections for heavy elements we have taken results for $h=0.1$, 0.075, and 0.05 and extrapolated to $h=0$. [*Note:* It has since been established that these small errors result from a minor error in programming; once this is corrected, the error in bound-state normalizations becomes completely negligible.]

ii. Continuum-State Wave Functions

The two parameters of a continuum-state solution which are easily compared with the analytical results of the pure Coulomb potential are normalization and phase. The accuracy of the normalizations calculated with $h=0.05$ for some typical partial waves in low-Z (26) and high-Z (82 and 84) elements are shown for various energies in Table I. The errors in the higher partial waves are nearly independent of energy (away from threshold) and Z, but increase rapidly with κ. These arise from the use of the same switching point $JJ=10$ from power series to numerical integration for all κ. Except at the highest energy (2 MeV), the errors are not sensitive to choices of h in the range from 0.05 to 0.1. The error in normalization also increases rapidly as the threshold energy is approached. This type of error is insensitive to h and may be connected with the very long periods of low-energy continuum states. In Table II we give the analogous errors in phase shifts, expressed both in radians and in percent.

TABLE I. Accuracy of continuum normalizations.

	Energy (MeV) \ κ	+1 (%)	+5 (%)	+9 (%)	+15 (%)
Low Z	0.354	0.01	0.02	0.4	4.0
	1.131	0.01	0.03	0.4	4.5
High Z	0.200	0.5	0.7	1.0	5.0
	0.354	0.1	0.2	0.5	4.5
	0.600	0.07	0.10	0.45	4.4
	1.131	0.03	0.05	0.4	4.2
	1.5	0.02	0.03	0.3	4.0
	2.0	0.2	0.2	0.4	4.2

TABLE II. Accuracy of phase shifts; errors in radians and percent.

	Energy (MeV) \ κ	+1	+5	+9	+15
Low Z	0.354	0.0004(2%)	0.0004(0.1%)	0.0005(0.1%)	0.0009(0.1%)
	1.131	0.0015(10%)	0.0016(0.5%)	0.0014(0.3%)	0.0014(0.2%)
High Z	0.200	0.0035(∞)	0.0054(0.3%)	0.0001(0.0%)	0.0021(0.1%)
	0.354	0.0030(1.5%)	0.0005(0.0%)	0.0019(0.1%)	0.0034(0.2%)
	0.600	0.0007(0.2%)	0.0001(0.0%)	0.0006(0.1%)	0.0012(0.1%)
	1.131	0.0011(0.3%)	0.0009(0.1%)	0.0011(0.1%)	0.0014(0.1%)

iii. Spherical Bessel Functions

The spherical Bessel functions j_l, $0 \leq l \leq 20$, were computed by the method of Corbató and Uretsky[14] and were checked against the NBS Tables[15] for selected values of l. The errors were less than one part in 10^6.

iv. Legendre Polynomials

The Legendre and associated Legendre functions were computed in double precision arithmetic directly from the polynomials given by Tallquist.[16] These were compared with double precision values computed from the recursion relations and found to agree to at least one part in 10^{12}. Rough comparisons can be made to the tables of Mursi[17] and NBS.[18]

B. The Main Program

i. Partial-Wave Integrals R_κ

The errors in an R_κ are due first to the errors of the component programs and second to the errors of the

[14] F. J. Corbató and J. L. Uretsky, J. Assoc. Comp. Mach. 6, 366 (1959).

[15] National Bureau of Standards, *Tables of Spherical Bessel Functions* (Columbia University Press, New York, 1947), Vols. I and II.

[16] H. J. Tallquist, Acta Soc. Sci. Fennicæ, Nova Series A, Tome II, No. 4 (1936); Tome II, No. 11 (1938).

[17] Z. Mursi, *Tables of Legendre Associated Functions* (E. and R. Schindler, Cairo, 1941).

[18] National Bureau of Standards, *Tables of Associated Legendre Functions* (Columbia University Press, New York, 1945).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Regions contributing to partial-wave integrals.

Z	Energy (MeV) \ κ	+1	+5	+9	+15
26	0.354	84	80	108	(does not contribute)
	1.131	58.5	63	78.5	89.5
50	0.354	41	45	50	(does not contribute)
92	0.208	19	28.5	35	(does not contribute)
84	0.354	20	26	31.5	(does not contribute)
	1.131	22	25	27	31
82	1.1368	21.5	23	27	28.5
	2.0	19.5	25	26	27
	2.754	22.5	24.5	25	25.5

integration procedures. The errors in integration arise from the choice of h and r_{max}, the upper limit of the integration.

Table III summarizes, for representative parameters, the distance in r which it was necessary to integrate to hopefully reduce the residual contribution to less than one part in 10^6. The needed distance increases slowly with increasing κ; in heavy elements it is nearly independent of energy and is determined by the decreasing exponential of the bound-state wave function. Since this exponential scales as Z the required distance greatly increases as Z decreases. In light elements this has the consequence that an energy dependence persists in the energy range of concern, and the rapid oscillations of a high-energy continuum-state work to reduce the needed distance.[19] Since such a stringent criterion was applied in the table, it is possible to obtain satisfactory results from the present calculation for a Z as low as 10 or 13, even with the restriction $r_{max} \leq 120$ required in the machine program.

The error in integration arising from the choice of h can be estimated by noting the change in the values of R_κ resulting from changes in h and subtracting out the portion of this which is due to change in component programs with h. As already noted, up to 1.5 MeV, the only component program which is sensitive to choices of h in the range 0.05–0.10 is the bound-state wave function. When the estimated error from this source is subtracted, the residual error to be attributed to error from h in integration is generally small and, with one exception,[20] completely independent of κ. The typical error begins to increase at higher energies: the difference between R_κ of $h=0.05$ and 0.10, still negligible at 662 keV, is about 0.4% at 1.131 MeV, both for $Z=26$ and 82, and 1.5% at 1.5 MeV.

TABLE IV. Number of κ's needed for accuracy to one part in 10^n.

E \ n	1	2	3	4	5	6
0.140	2	3	4	4	5	5
0.200	3	4	5	6	7	8
0.279	3	4	6	7	8	9
0.354	4	5	7	9	10	12
0.400	4	6	8	9	11	13
0.662	5	8	10	13	14	
1.131	7	12	15			
1.1368	8	13	16			
2.0	11	≈17				
2.754	14					

[19] It should be noted that the distance required for very high accuracy has a different dependence on parameters from the distance which characterizes the main region of the integrand. It is the latter distance which is discussed in the various analytic approximations.

ii. Number of κ's Contributing to Cross Section

The number of κ's needed to determine the total cross section to any desired accuracy increases with energy, but it is almost independent of Z. One or two more κ's are needed to obtain the same accuracy for $Z=26$ as for $Z=84$. In Table IV we summarize as a function of energy (for heavy elements) the number of κ's needed to obtain an accuracy of one part in 10^n. The restriction to $|\kappa|<20$ results in a limitation to energies below about 2 MeV if it is desired to obtain the total cross section to 1%.

TABLE V. Estimates of total error in numerical calculation.

	Energy	Error in σ_{total} (%)
Low Z	0.354	0.2
	1.131	0.6
High Z	0.200	1.1
	0.354	0.5
	0.600	0.5
	1.131	0.8
	1.368	0.8
	2.00	2.0

iii. Estimate of Total Error in Cross Sections

We assume that a fit is made to the results for different h and extrapolated to $h=0$, and we assume that, in agreement with our analysis, this renders errors associated with h (bound-state normalization, integration, etc.) negligible, i.e. 0.1%. Then the error in R_κ is dominated by the error in the continuum wave function. From this estimate of the error in each R_κ, and from the information just presented on the contribution of each κ to the cross section, we can get an estimate of the total error of the cross section. Some rough estimates of this kind are summarized in Table V. The main conclusion is that the present calculation obtains total cross sections accurate to 0.8% in the energy range from 300 to 1400 keV; the accuracy decreases rapidly toward threshold and at high energies.

V. TOTAL CROSS SECTIONS

Theoretical predictions for the total K-shell photoelectric-effect cross section have been based on four

[20] The exception is for $\kappa=1$, for which R_κ changes by as much as 0.5% between $h=0.05$ and 0.10.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VI. Total cross sections for $Z=13$, 26, 50, 84, in barns.

Z	13	26		50		84	
E (MeV)	This work	This work	Hulme *et al.*	This work	Hulme *et al.*	This work	Hulme *et al.*
0.3543	0.01543	0.3885	0.39	7.06	7.1	60.6	60.2
0.662	0.00289	0.0754	···	1.455	···	13.87	···
1.131	0.00091	0.0237	0.023	0.462	0.46	4.54	4.61
1.5	0.00054	0.0140	···	0.271	···	2.65	···
2.0	0.00033	0.0086	···	0.164	···	1.57	···

TABLE VIII. Comparison of total cross sections (in barns) for $Z=82$, from the NBS tables, an extrapolation of Pratt, and the present calculation.

E (MeV)	NBS	Extrapolation	Present calculation
1.131	4.37	4.07	4.10
1.332	3.24	3.03	2.99
1.5	2.65	2.49	2.39
2.0	1.8	1.56	1.42

main calculations: (1) The exact nonrelativistic solution of the problem, usually associated with Stobbe,[21] (2) Sauter's relativistic calculation,[6] valid to lowest order in $Z\alpha/\beta$, (3) numerical calculations of Hulme *et al.*[22] for a few selected energies and elements, (4) the high-energy limit of the relativistic problem for arbitrary Z as given by Hall.[23] Extrapolation procedures, based on these four results, can be used to estimate the cross section for arbitrary Z and energy; tables of these predictions have been given by Grodstein[24] (NBS tables). Experimental results have generally yielded satisfactory agreement with these tables.

Since 1955 there has been substantial improvement in all four of these calculations, and one of them has actually been found to contain an error. Namely, although Hall's expression for the high-energy limit in the form of a double integral is correct, the analytic formula with which he approximated the integral is not, and overestimates the cross section by a factor of two in heavy elements. The exact high-energy limit was obtained numerically by Pratt[7] and later confirmed by Hall[25]; as already noted, Eq. (2.7) gives a fairly good analytic representation. It is then at first hard to understand why the extrapolation between Sauter and Hall is in good agreement with Hulme. The work of Gavrila[26] and Nagel,[27] who extended Sauter's result to next order in $(Z\alpha/\beta)$, explains the puzzle. Hall's error essentially involved omission of the factor $(-4\pi a/15)$ in Eq. (2.7), but the energy dependence of this term is such that it is large only at very high energies. [Recently, a complicated analytical expression for still another order in $Z\alpha/\beta$ has been given by Gorshkov and Mikhailov.[28]]

K-shell photoeffect at threshold has been calculated for the relativistic problem, both analytically and numerically, by Nagel and Olsson.[5] In the low-energy region screening effects are expected to be important, however, and this has recently been examined in the nonrelativistic problem by Cooper.[29] Since, as already noted, the numerical techniques of the present paper require modification at low energies, we will not discuss this energy region further.

The numerical calculations of Hulme *et al.* provided the reference points which Sauter-Stobbe and Sauter-Hall extrapolations are forced to fit. The accuracy of Grodstein's predictions for intermediate energies rests almost entirely on the accuracy of these numerical results obtained for 0.3543 and 1.131 MeV, estimated at 4% for heavy elements and 8% for lighter elements. Since this is the main energy region of experimental interest, it is important to verify and extend these

TABLE VII. Total cross sections for $Z=82$, 92, in barns.

	$Z=82$				$Z=92$		
E (MeV)	This work	HNO	NBS	E (MeV)	This work	HNO	NBS
0.120	···	921	950				
0.200	239	241	236	0.132	···	1026	1100
0.300	84.0	83.6	80.2	0.140	···	887	942
0.400	40.7	40.5	40.4	0.208	319	324	334
0.500	23.7	23.7	23.1	0.279	155	154	155
0.600	15.6	15.5	15.3	0.412	59.9	59.5	59.8
0.662	12.5	···	12.4	0.662	20.4	20.2	20.7
0.900	6.49	···	6.68	0.900	10.7	···	11.2
1.131	4.10	···	4.37	1.131	6.78	···	7.38
1.332	2.99	···	3.24	1.332	4.93	···	5.5
1.5	2.39	···	2.65	1.5	3.95	···	4.5
2.0	1.42	···	1.8	2.0	2.33	···	2.9

TABLE IX. Comparison with experimental results (see Refs. 32–34) for K-shell total cross sections (in barns).

Experimenter	Element	Energy (keV)	Experimental result	Present theory
Seeman	Pb	511	23.4 ±0.7	22.5
Missoni	Au	662	10.2 ±0.3	10.5
Hultberg and Stockendal	U	1173	7.2 ±0.5	6.32
Hultberg and Stockendal	U	1332	5.4 ±0.3	4.93
Bleeker, Goudsmit and De Vries	Pb	1332	3.24±0.13	2.99

[21] M. Stobbe, Ann. Physik 7, 661 (1930).

[22] H. R. Hulme, J. McDougall, R. A. Buckingham, and R. H. Fowler, Proc. Roy. Soc. (London) A149, 131 (1935).

[23] H. Hall, Rev. Mod. Phys. 8, 358 (1936).

[24] G. W. Grodstein, *U. S. Department of Commerce, National Bureau of Standards Circular 583* (U. S. Government Printing Office, Washington, D. C., 1957); see also R. T. McGinnies, *NBS Supplement to Circular 583* (1959).

[25] H. Hall, University of California, Lawrence Radiation Laboratory Report UCRL 5947-T (unpublished).

[26] M. Gavrila, Phys. Rev. 113, 514 (1959).

[27] B. Nagel, Arkiv Fysik 18, 1 (1960).

[28] V. G. Gorshkov and A. I. Mikhailov, Zh. Eksperim. i Teor. Fiz. 43, 991 (1962) [English transl.: Soviet Phys.—JETP 16, 701 (1963)].

[29] J. W. Cooper, Phys. Rev. 128, 681 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE X. K-shell angular distributions for $Z=13, 26, 50, 84$ in barns/steradian (when multiplied by indicated scale factor and by 1490).

Z	13		26		50			84		
E (keV)	354	662	354	662	354	662	1131	354	662	1131
θ Scale	10^{-6}	10^{-6}	10^{-4}	10^{-5}	10^{-3}	10^{-3}	10^{-4}	10^{-2}	10^{-3}	10^{-3}
0			0.0028		0.0155	0.033	0.4		0.79	0.96
5	1.0	0.5	0.19	1.4	0.280	0.22	2.0	0.14	1.91	1.98
10	2.8	1.4	0.62	3.6	0.92	0.59	4.3	0.43	3.93	3.40
15	4.1	1.7	1.04	4.4	1.58	0.79	4.3	0.78	5.20	3.40
20	4.9	1.6	1.27	4.1	2.00	0.76	2.8	1.06	5.50	2.47
25	5.3	1.3	1.28	3.3	2.10	0.598	1.6	1.23	4.99	1.62
30	4.9	0.9	1.15	2.3	1.95	0.435	1.02	1.26	3.98	1.07
35	3.9	0.58	0.95	1.50	1.67	0.306	0.67	1.20	2.98	0.71
40	2.9	0.37	0.746	1.00	1.36	0.211	0.42	1.08	2.22	0.47
45	2.2	0.26	0.567	0.71	1.07	0.144	0.27	0.93	1.64	0.33
50	1.65	0.17	0.422	0.48	0.820	0.098	0.18	0.78	1.19	0.23
55	1.20	0.11	0.311	0.30	0.621	0.069	0.13	0.64	0.89	0.168
60	0.86	0.08	0.228	0.21	0.468	0.049	0.095	0.51	0.66	0.125
65	0.63	0.06	0.168	0.16	0.351	0.035	0.069	0.41	0.49	0.097
70	0.46	0.04	0.123	0.116	0.264	0.026	0.051	0.33	0.37	0.076
75	0.34	0.028	0.091	0.077	0.199	0.019	0.039	0.26	0.29	0.058
80	0.26	0.020	0.067	0.056	0.150	0.0142	0.031	0.21	0.23	0.046
85	0.19	0.016	0.050		0.113	0.0109	0.024	0.17	0.17	0.038
90	0.140	0.012	0.037		0.086	0.0083	0.019	0.13	0.14	0.032
95	0.101		0.028		0.066	0.0064				
100	0.076		0.021		0.051	0.0050				
105	0.057		0.0158		0.039	0.0040				
110	0.042		0.0120		0.030	0.0032				
115	0.033		0.0091		0.024	0.0026				
120	0.027		0.0069		0.019	0.0020				
125	0.020		0.0052		0.015	0.0017				

Z	13			26			50		84	
E (keV)	1131	1500	2000	1131	1500	2000	1500	2000	1500	2000
θ Scale	10^{-6}	10^{-6}	10^{-6}	10^{-5}	10^{-5}	10^{-5}	10^{-4}	10^{-4}	10^{-3}	10^{-3}
0				0.06	0.05	0.034	0.4	0.36	1.0	1.0
5	0.42	0.4	0.3	1.0	1.0	0.9	2.0	1.8	2.0	1.9
10	0.95	0.79	0.6	2.4	2.0	1.6	3.7	3.0	3.0	2.6
15	0.90	0.63	0.43	2.36	1.60	1.04	2.9	1.87	2.43	1.6
20	0.58	0.31	0.16	1.50	0.80	0.41	1.5	0.76	1.41	0.72
25	0.33	0.15	0.08	0.83	0.39	0.19	0.78	0.37	0.81	0.39
30	0.195	0.09	0.05	0.50	0.24	0.126	0.49	0.24	0.51	0.24
35	0.119	0.061	0.030	0.33	0.17	0.079	0.32	0.15	0.32	0.146
40	0.075	0.034	0.015	0.205	0.091	0.040	0.19	0.086	0.21	0.099
45	0.048	0.019	0.100	0.122	0.053	0.024	0.119	0.056	0.15	0.071
50	0.031	0.014	0.008	0.081		0.014	0.085	0.043	0.105	0.047
55	0.020	0.011		0.060			0.064	0.030	0.075	0.034
60	0.015	0.007					0.045	0.021	0.058	0.030
65	0.012						0.032	0.015	0.047	0.023
70	0.008						0.025	0.012	0.036	0.016

calculated points, and this is the purpose of the present section.

While the present work was in progress a new numerical calculation was reported by Hultberg, Nagel, and Olsson (HNO).[2] Like the work of Hulme *et al.*, the method used applies only to the pure Coulomb potential; HNO treated two heavy elements in the energy range 0.120–0.662 MeV. Moderately good agreement was obtained with the values derived from the NBS tables, confirming the low-energy high-Z point of Hulme *et al.*[30]

The K-shell total cross sections obtained in the present numerical calculations are summarized in Tables VI and VII. The Z's of Table VI corresponds to those of

[30] It should also be noted that in principle, another series of photoeffect total cross sections is available in the numerical calculations on the coherent scattering of photons from K-shell electrons carried out by Brown and collaborators: G. E. Brown, R. E. Peierls, and J. B. Woodward, Proc. Roy. Soc. (London) **A227**, 51 (1954); S. Brenner, G. E. Brown, and J. B. Woodward, *ibid.* **A227**, 59 (1954); G. E. Brown and D. F. Mayers, *ibid.* **A234**, 387 (1955); **A242**, 89 (1957). As the authors note, the photoeffect cross sections can be obtained from the imaginary part of the scattering amplitude in the forward direction for the coherent process, and results are presented for Hg at photon energies of 0.32, 0.64, 1.28, and 2.56 (all in units of mc^2). A similar calculation was later reported by H. Cornille and M. Chapdelaine, Nuovo Cimento **14**, 1386 (1959) for photons of energy 5.12 mc^2 on Hg. Unfortunately, results of these papers pertaining to the photoeffect are not presented as cross sections. We do not wish to now enter into a detailed discussion, but if our interpretation of these results is correct they are in fair accord with our numerical work and the highest energy results show the suppression from previous values reported here.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE XI. K-shell angular distributions for $Z=82$, 92 in barns/steradian (when multiplied by indicated scale factor and by 1490).

Z			82					92		
E (keV)	200	300	400	500	600	662	208	279	412	662
θ \ Scale	10^{-2}	10^{-2}	10^{-2}	10^{-2}	10^{-3}	10^{-3}	10^{-2}	10^{-2}	10^{-2}	10^{-2}
0	0.0006		0.0289	0.0414	0.523	0.577	0.00030	0.0147	0.052	0.104
5	0.132	0.2	0.135	0.143	1.51	1.55	0.122	0.139	0.175	0.228
10	0.499	0.5	0.397	0.376	3.63	3.56	0.47	0.479	0.484	0.489
15	1.02	0.8	0.688	0.600	5.33	4.97	0.98	0.94	0.85	0.690
20	1.61	1.1	0.900	0.717	5.82	5.16	1.57	1.42	1.13	0.740
25	2.16	1.36	0.987	0.717	5.34	4.50	2.16	1.83	1.27	0.673
30	2.59	1.53	0.964	0.642	4.43	3.58	2.69	2.10	1.28	0.577
35	2.88	1.54	0.871	0.536	3.47	2.71	3.10	2.23	1.19	0.437
40	3.02	1.44	0.746	0.428	2.63	2.00	3.37	2.23	1.05	0.332
45	3.02	1.28	0.616	0.333	1.96	1.46	3.50	2.12	0.89	0.250
50	2.91	1.10	0.497	0.255	1.45	1.06	3.50	1.96	0.74	0.187
55	2.72	0.93	0.395	0.194	1.07	0.78	3.40	1.76	0.60	0.140
60	2.48	0.78	0.311	0.148	0.80	0.576	3.22	1.54	0.49	0.106
65	2.22	0.64	0.244	0.112	0.60	0.431	2.98	1.33	0.39	0.080
70	1.96	0.53	0.190	0.086	0.455	0.327	2.72	1.14	0.310	0.062
75	1.71	0.43	0.149	0.066	0.348	0.250	2.45	0.96	0.247	0.048
80	1.47	0.35	0.117	0.051	0.269	0.193	2.18	0.81	0.197	0.038
85	1.26	0.28	0.092	0.040	0.209	0.151	1.93	0.68	0.158	0.030
90	1.07	0.23	0.073	0.031	0.166	0.120	1.69	0.57	0.127	0.024
95	0.91	0.183	0.058	0.025	0.133	0.097	1.47	0.47	0.103	0.0196
100	0.766	0.148	0.047	0.0200	0.108	0.079	1.27	0.39	0.083	0.0161
105	0.644	0.121	0.038	0.0163	0.088	0.064	1.10	0.33	0.068	0.0134
110	0.54	0.101	0.031	0.0134	0.073	0.054	0.94	0.27	0.057	0.0113
115	0.45	0.083	0.026	0.0111	0.062	0.046	0.81	0.230	0.048	0.0098
120	0.38	0.068	0.021	0.0095	0.053	0.039	0.69	0.195	0.041	0.0085
125	0.31	0.056	0.018	0.0082	0.046	0.034	0.59	0.166	0.036	0.0075

Z			82					92		
E (keV)	900	1131	1332	1500	2000	900	1131	1332	1500	2000
θ \ Scale	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}	10^{-3}
0	0.75	0.84	0.90	0.90	0.87	1.37	1.56	1.68	1.7	1.68
5	1.71	1.79	1.84	1.8	1.75	2.62	2.82	2.9	2.9	2.9
10	3.34	3.12	2.94	2.76	2.3	4.81	4.63	4.42	4.2	3.6
15	3.89	3.11	2.57	2.22	1.4	5.68	4.67	3.93	3.4	2.25
20	3.32	2.24	1.63	1.27	0.65	5.04	3.51	2.59	2.04	1.06
25	2.46	1.46	0.98	0.73	0.35	3.88	2.39	1.63	1.22	0.59
30	1.73	0.96	0.62	0.46	0.22	2.83	1.61	1.05	0.77	0.36
35	1.20	0.64	0.402	0.290	0.131	2.01	1.08	0.68	0.49	0.217
40	0.82	0.423	0.264	0.189	0.088	1.42	0.74	0.46	0.33	0.153
45	0.57	0.291	0.183	0.133	0.063	1.01	0.52	0.33	0.236	0.111
50	0.41	0.207	0.129	0.093	0.042	0.73	0.37	0.231	0.166	0.072
55	0.294	0.148	0.092	0.066	0.030	0.54	0.27	0.166	0.118	0.053
60	0.216	0.110	0.070	0.051	0.026	0.40	0.202	0.129	0.093	0.048
65	0.164	0.085	0.056	0.042	0.020	0.31	0.160	0.104	0.077	0.037
70	0.126	0.066	0.042	0.031	0.014	0.24	0.125	0.080	0.059	0.025
75	0.096	0.051				0.185	0.096	0.061		
80	0.075	0.040				0.146	0.077	0.051		
85	0.061	0.033				0.119	0.065	0.044		
90	0.050	0.028				0.098	0.054	0.035		

Hulme *et al.* and those of Table VII to HNO; in each case the appropriate comparisons with earlier work are made. The first striking feature is the excellent agreement, far better than the accuracy which had been claimed, with the calculations of Hulme *et al.* This means that interpolations, such as Grodstein's, based on the Hulme values would be expected to be good and need no major revision in this energy range (however, see below). The errors which remain are of the same magnitude as screening effects; such effects must be included before the change in the cross section from the Hulme values is significant. Agreement with the HNO results is also excellent, and within the estimated error of the present calculation. The first result of this section is then the establishment of the accuracy of the values of Hulme *et al.*

In Table VII we have also given the corresponding cross section as obtained in the NBS tables.[31] The com-

[31] Since the NBS values are for total absorption cross sections, the contribution from higher shells is removed by dividing out the simple multiplicative factor which Grodstein used to put it in. Another smaller correction is also needed to obtain a K-shell cross section without screening.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

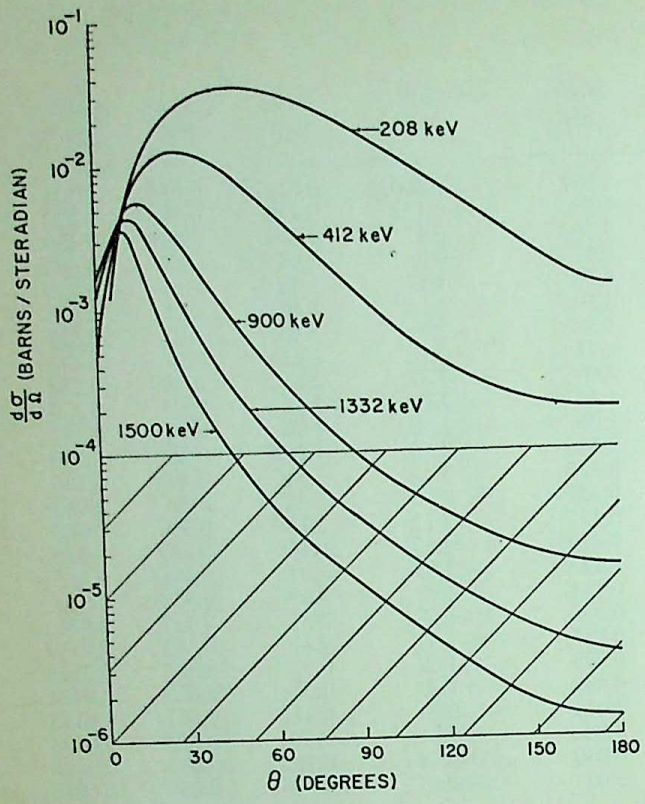


FIG. 1. K-shell angular distributions for U in barns/steradian. The small cross sections in the shaded region of the graph are of low accuracy.

parison is rather surprising; even in the 1-MeV region, the present results are 5–10% lower than corresponding interpolations from the NBS tables. It is hard to understand this, since the present results agree with Hulme *et al.* and the NBS tables are based on Hulme *et al.* By 2 MeV, the difference has reached 25%. Further proof that this difference is real is provided by a new calculation[32] which HNO have undertaken to check our results; it appears that they have obtained complete agreement with our cross section for $Z=92$ at 1.332 MeV.

A comparison with the extrapolation formula of Pratt[7] is given in Table VIII for $Z=82$. This formula, which is exact at very high energies and exact for light elements at all energies, is in good agreement with the Hulme values and so also with the present calculation in the 1-MeV range. However Table VIII suggests that the very close agreement at 1 MeV is misleading, since for the higher energy of 2 MeV the two differ by 10%.

The recent experimental measurements[33–36] of the total cross section from the K shell are given in Table IX.[37] These results were all reported as in good accord with the NBS tables, which again says that the experimental results above 1 MeV are higher than the theoretical values of the present calculation.

The second result of this section is then the finding of significantly lower cross sections in the energy region above 1 MeV than are given in the NBS tables or in two recent experiments. We have previously outlined the evidence for the accuracy of the present calculations. If these arguments are correct we see no theoretical explanation for the discrepancy. It is true that screening effects will modify the pure Coulomb calculations, but for the K shell such corrections are expected to be only 1–2%, and in the wrong direction.

TABLE XII. Comparison of angular distributions with HNO results, normalized to agree with this work at the underlined angles.

E(keV)	208		279		412		662	
θ	This work	HNO	This work	HNO	This work	HNO	This work	HNO
0			0.01		0.05	0.05	0.10	0.13
5	0.12	0.13	0.14	0.17	0.18	0.18	0.23	0.27
10	0.47	0.53	0.48	0.51	0.48	0.48	0.49	0.49
15	0.98	0.97	0.94	0.94	0.85	0.77	0.69	0.68
20	1.57	1.50	1.42	1.38	1.13	1.14	0.74	0.74
25	2.16	2.17	1.83	1.83	1.27	1.27	0.67	0.67
30	2.69	2.73	2.10	2.10	1.28	1.28	0.56	0.56
35	3.10	3.13	2.23	2.23	1.19	1.18	0.44	0.44
40	3.37	3.37	2.23	2.21	1.05	1.04	0.33	0.33
45	3.50	3.47	2.12	2.10	0.89	0.88	0.25	0.25
50	3.50	3.50	1.96	1.95	0.74	0.73	0.19	0.19
55	3.40	3.43	1.76	1.76	0.60	0.61	0.14	0.14
60	3.22	3.23	1.54	1.53	0.49	0.49	0.11	0.11
65	2.98	3.00	1.33	1.34	0.39	0.39	0.08	0.08
70	2.72	2.77	1.14	1.13	0.31	0.30	0.06	0.06
75	2.45	2.47	0.96	0.96	0.25	0.24	0.05	0.05
80	2.18	2.20	0.81	0.81	0.20	0.20	0.04	0.04
85	1.93	1.93	0.68	0.68	0.16	0.16	0.03	0.03
90	1.69	1.70	0.57	0.57	0.13	0.12		
95	1.47	1.47	0.47	0.47	0.10	0.10		
100	1.27	1.27	0.39	0.40	0.08	0.09		
105	1.10	1.10	0.33	0.34	0.07	0.07		
110	0.94	0.93	0.27	0.28	0.06	0.06		
115	0.81	0.80	0.23	0.23	0.05	0.05		
120	0.69	0.67	0.20	0.19				
125	0.59	0.50	0.17	0.17				

[32] S. Hultberg (private communication). We should like to thank Dr. Hultberg for undertaking this calculation and for communicating its results prior to publication.

[33] K. W. Seeman, Bull. Am. Phys. Soc. 1, 198 (1956).

[34] S. Hultberg and R. Stockendal, Arkiv Fysik 15, 355 (1959).

[35] E. J. Bleeker, P. F. A. Goudsmit, and C. DeVries, Nucl. Phys. 29, 452 (1962).

[36] G. Missoni (to be published).

[37] One should also note the measurement by Bleeker, Goudsmit, and DeVries of the ratio $\sigma(2.754)/\sigma(1.368)$ in Pb. A similar ratio was obtained earlier by G. R. Bishop, C. H. Collie, H. Halban, A. Hedgran, K. Siegbahn, S. duToit and R. Wilson, Phys. Rev. 80, 211 (1950); an early result for $\sigma(2.62)$ in Pb is given by G. D. Latyshev, Rev. Mod. Phys. 19, 132 (1947). See also the total photoelectric absorption cross-section measurements of W. F. Titus, Phys. Rev. 115, 351 (1959), at 662 keV in a series of elements, and of B. I. Deutch and F. R. Metzger, *ibid.* 122, 848 (1961) at 279 keV in Tl. Particularly interesting for this paper are total absorption measurements at 2.62 MeV in various elements by I. E. Dayton, Phys. Rev. 89, 544 (1953) and W. F. Titus (to be published), which lie appreciably below the earlier theories.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE XIII. Ratio of experiment to theory for angular distributions in U, both distributions normalized to unity at the underlined angles.

Angle \ Experimenter / Energy (keV)	Sujkowski 279	Hultberg 412	Hultberg 662	Hultberg 1332
0°	6.2	2.2	0.89	0.71
15°	0.95	0.84	1.00	1.00
30°	1.00	1.00	1.04	1.28
45°	1.05	1.10	1.32	0.87
60°	0.94	1.33	1.46	0.91
75°	0.92	1.58	1.68	0.67
90°	0.98	1.90	1.69	0.00

VI. ANGULAR DISTRIBUTIONS

Until quite recently the only theoretical result for the angular distribution of relativistic K-shell photoeffect was that of Sauter,[6] Eq. (2.5), and there was no experimental information. When corrections of relative $0(a)$ were computed by Banerjee,[38] Gavrila,[26] and Nagel[27] it was found that they did not significantly change the predicted angular distributions, including the prediction of a vanishing cross section in the forward direction. On the other hand, the experiments of Hultberg[39–41] indicated some striking deviations from the Sauter distribution: A nonvanishing forward cross section, a shifting of the maximum in the cross section toward *larger* angles, and a tendency of the cross section to hold up at large angles. We shall return to the theoretical discussion of these matters following the presentation of our numerical calculations and their comparison with HNO and experiment.

The K-shell angular distributions obtained in the present numerical calculations are summarized in Tables X and XI. We present these results in some detail, despite their limited accuracy, since they do display the significant deviations from the Sauter distribution and so provide more nearly correct predictions. The very small large-angle cross sections are not tabulated. For qualitative purposes we also show in Fig. 1 the $Z=92$ distributions for several energies. Forward scattering increases with increasing energy, the maximum moves in, and backward scattering remains finite but decreases with increasing energy.

For the lower energies we may again compare these results with those of HNO which become available (although they have not been published) in the course of the present calculation. The comparison for U is given in Table XII, where the HNO values, which are normalized to unity at the angle of maximum emission, are read from their graphs and normalized to agree with our absolute results at the underlined angles. The two calculations are in good agreement.

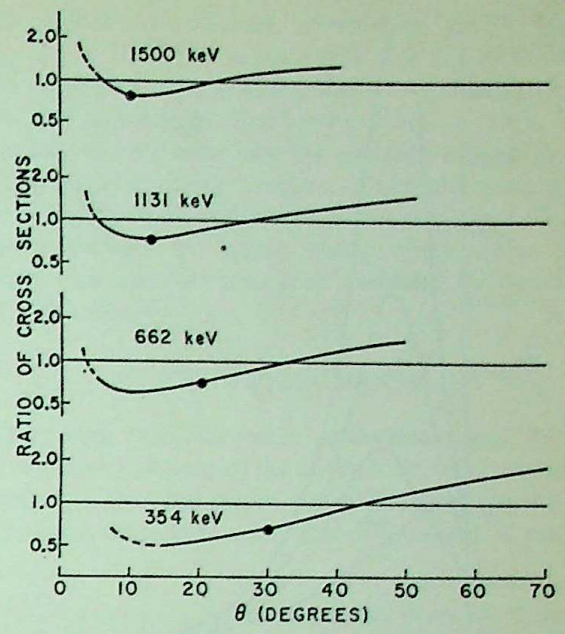


FIG. 2. Comparison with the Sauter angular distribution from $Z=84$ for a series of energies. The ratio R of the present distribution to the Sauter distribution is plotted, where both distributions have been normalized to their respective total cross sections. The approximate angles of maximum emission are marked.

Experimental results for K-shell angular distributions have been obtained for uranium by Hultberg[39] (412, 662, and 1332 keV) and by Sujkowski[40] (279 keV); no other true K-shell distributions are known to us, excepting the early work of Hultberg.[37–38] In Table XIII we give, for a few selected angles, the ratio of experiment to theory. The experimental results are available at quite fine angular intervals, and the 279- and 662-keV cases have elsewhere[2,42] been compared with the HNO predictions in graphical form. Except at forward angles the 279-keV result of Sujkowski seems to be in good agreement with theory, and it should be noted that at this low energy the forward cross section is very small. The higher energy data of Hultberg shows more marked deviations from theory; now that fairly good theoretical estimates are available, further experimental data in this energy region would be desirable. Note that at higher energies theory predicts a *larger* forward cross section than was observed. [Angular distributions for Au at 412 keV in the range 20°–90° have now been obtained by Bergkrist and Hultberg[43] in excellent agreement with theory.]

Let us next compare these new theoretical results with the much-used Sauter distribution. This has been done for a series of energies in Fig. 2. What is plotted is the ratio R of the present distribution to the Sauter distribution, with both distributions normalized to their respective total cross section $[(d\sigma/d\Omega)/\sigma]$. Thus, the

[38] H. Banerjee, Nuovo Cimento **10**, 863 (1958).
[39] A. Hedgran and S. Hultberg, Phys. Rev. **94**, 498 (1954).
[40] S. Hultberg, Arkiv Fysik **9**, 245 (1955).
[41] S. Hultberg, Arkiv Fysik **15**, 307 (1959).
[42] Z. Sujkowski, Arkiv Fysik **20**, 269 (1961).
[43] K. Bergkvist and S. Hultberg, Arkiv Fysik (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

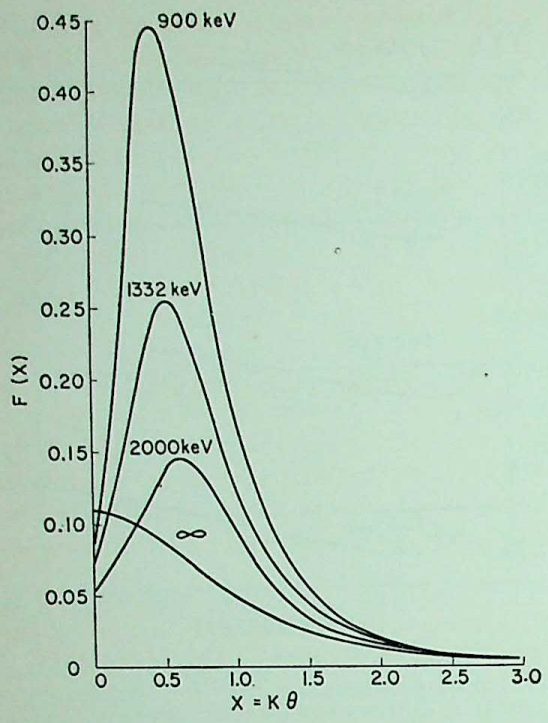


FIG. 3. High-energy behavior of angular distribution. $F(x) = (4e^2a^5\epsilon)^{-1}(d\sigma/d\Omega)$ is plotted against $x=k\theta$ for $Z=82$ at three energies and compared with its form in the high-energy limit.

straight line $R=1$ corresponds to the Sauter distribution; an angular region in which $R=$constant is a region in which the distribution locally has the Sauter shape. The only case shown is for $Z=84$. For $Z=26$ the curves essentially follow the straight lines $R=1$ and even for $Z=50$ the deviations are barely significant within the errors of the calculation This was why, as was elucidated by later theoretical work, it was possible to get along with a Sauter distribution for many years—the shape was much better than the normalization. Roughly speaking, the error in the Sauter total cross section is $0(\pi a)$ (from omission of the characteristic $e^{-\pi a}$ factor) but the significant errors in the Sauter distribution are only $0(a^2)$. But Fig. 2 does show that in the heaviest elements the deviations from the Sauter distribution become significant. The rise in R at small angles occurs because the Sauter distribution, unlike an exact calculation, vanishes in the forward direction. If this were the only region which deviated from Sauter form the curve would elsewhere behave as $R=$constant. Instead, at larger angles it falls off more slowly than the Sauter form, and indeed there is no sizeable angular region of Sauter shape. These qualitative features were first remarked in the experiments. Finally, the angle $\theta_{\max}$ of maximum emission, roughly indicated on the graphs, occurs in the region in which R is rising, so the cross section is staying up in comparison to Sauter, and so $\theta_{\max}$ is being shifted toward larger angles in comparison to the Sauter distribution. This too was first noted in the experiments.

The relative $0(a)$ corrections, mentioned earlier, do not shed much light on these properties—they appear to be $0(a^2)$ effects. This arises, as was noted by Nagel,[27] both because at intermediate energies the $0(a)$ term is not large and because its distribution is similar to the Sauter form. There is a tendency for a maximum at larger angles and for a larger cross section at back angles, but especially for the latter effect the increase is insufficient. Nagel also pointed out that the reason the $0(a)$ distribution vanishes in the forward direction is that it arises as a cross term with the Sauter matrix element. The lowest order nonvanishing cross section in the forward direction $I(0)$ is relative $0(a^2)$, and this Nagel obtained.[43a] It has been conventional to express this in terms of the ratio $\kappa \equiv I(0)/I(\theta_{\max})$. From Table XI we would estimate that at 1332 keV, this ratio is 0.30 for Pb and 0.38 for U. These compare with the theoretical values 0.27 obtained for Pb by Sauter and Wüster[44] in a numerical calculation and 0.42 obtained by Nagel for Pb from the relative $0(a^2)$ term above. Hultberg's experimental value for κ in U at 1332 keV was 0.245, which again reflects the fact that the experimental forward cross section at the higher energies lies below theory.[45] (Some new measurements of κ have recently been reported by Rimskii-Korsakov *et al.*[46])

It is finally of some interest to discuss the behavior of the K-shell angular distribution in the high-energy limit. At high energies, the significant angles are $0(1/k)$, and in analytic work it is appropriate to expand in $1/k$ and θ, but for arbitrary $x \equiv k\theta$. Neglecting $0(a^2)$, Nagel[27] writes the distribution in this limit as

$$\sim \frac{x^2}{(1+x^2)^3}\left[1-a\frac{\pi}{2}\frac{1}{(1+x^2)^{1/2}}\right], \qquad (6.1)$$

including the $0(a)$ correction to the Sauter distribution which has its maximum at $x=1/\sqrt{2}$. Nagel (and also Mork and Olsen[47]) find that in this limit forward scattering is characterized by $\kappa=5.85a^2$ to lowest nonvanishing order in a. Mork and Olsen also obtained an exact expression for the forward scattering in the high-

[43a] This result was also independently obtained by Arne Reitan, Physica Norvegica 1, 113 (1961).

[44] F. Sauter and H. O. Wüster, Z. Physik 141, 83 (1955). Actually these authors obtained $I(0°)$ numerically and took $I(\theta_{\max})$ from the usual Sauter distribution, a procedure which certainly becomes invalid at high energies as will be seen shortly.

[45] Nagel (Ref. 27) compares the Sauter-Wüster 0.27 value with Hultberg's 0.245, calling the agreement rather good, apparently assuming that for large Z, κ does not change appreciably with Z. The present results do not support this assumption, and we conclude there is a significant difference between theory and experiment.

[46] K. K. Aglinstev, V. V. Mittrofanov, A. A. Rimskii-Korsakov, and V. V. Smirnov, Bull. Acad. Sci. USSR 25, 1146 (1962); A. A. Rimski-Korsakov and V. V. Smirnov, Zh. Eksperim. i Teor. Fiz. 42, 67 (1962) [English transl.: Soviet Phys.—JETP 15, 47 (1962)]; and Bull. Acad. Sci. USSR 26, 1180 (1963).

[47] K. Mork and H. Olsen, Proceedings of the Physics Seminar in Trondheim, No. 5, 1960 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

energy limit; such an expression was also later obtained by Weber and Mullin,[48] who graph $F(0)$ against Z, where

$$F(\theta)=\frac{1}{4e^2a^5\epsilon}\frac{d\sigma}{d\Omega} \tag{6.2}$$

in dimensionless units. In the high-energy limit, F is independent of energy. For Pb we would compute $\kappa=2.1$, an absurd result ($\kappa\leq 1$) which could be interpreted as a gross failure of the expansion of $F(0)$ in a. In fact, the error lies in the supposition that the position of the maximum is given by the Sauter maximum, whereas numerical calculations of Boyer[49] and Nagel[50] show that in heavy elements the maximum occurs for forward emission and the distribution decreases monotonically. Using Nagel's calculation[50] we find that the ratio of forward emission to emission at the Sauter angle is 1.6. In Fig. 3 we plot $F(x)$ for $Z=82$ for three of the energies of our numerical calculation and also show the high-energy limit from Nagel's work. This again makes very clear how far removed 2 MeV is from the asymptotic energy region. Even the manner in which the minimum at forward angles gets filled remains to be calculated by higher energy work. (The high-energy limit has now also been discussed by Gorshkov and Mikhailov.[51])

ACKNOWLEDGMENTS

This work was suggested and encouraged by Dr. Harvey Hall, to whom all the authors are much indebted. Thanks are also due to Dr. Solve Hultberg, for helpful discussions, correspondence, and prepublication results.

[48] T. A. Weber and C. J. Mullin, Phys. Rev. 126, 615 (1962).
[49] R. H. Boyer, Ph.D. thesis, Oxford University, 1957 (unpublished).
[50] B. Nagel, Arkiv Fysik 24, 151 (1963).
[51] V. G. Gorshkov and A. I. Mikhailov, Zh. Eksperim. i Teor. Fiz. 44, 2142 (1963) [English transl.: Soviet Phys.—JETP (to be published)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Polarization Correlations in Atomic Photoeffect*

R. H. PRATT,†
Institute of Theoretical Physics, Department of Physics, Stanford University, Stanford, California
AND
RICHARD D. LEVEE,‡ ROBERT L. PEXTON, AND WALTER ARON
Lawrence Radiation Laboratory, University of California, Livermore, California
(Received 13 September 1963)

The seven possible polarization correlations between incident photon and ejected electron in K-shell photoeffect have been obtained with numerical methods for the range of elements from $Z=13$ to $Z=92$ and the range of energies from 200 keV to 2 MeV. The photoeffect can serve as a polarizer of electrons, a transmitter of polarization from photons to electrons, or an analyzer of polarized radiation. In heavy elements at suitable angles of emission all these correlations can be large. Previous work on the correlations is discussed.

I. INTRODUCTION

THE atomic photoelectric effect can serve as an analyzer of polarized radiation, a transmitter of polarization from photons to electrons, or a polarizer of electrons. The preferential ejection of photoelectrons in the plane defined by the momentum $\mathbf{k}$ and polarization vector $\mathbf{e}$ of linearly polarized photons has long been known.[1] In recent years the high degree of correlation between circularly polarized photons and longitudinally polarized electrons has been discussed.[2–6] The photoeffect can also serve as a polarizer, producing transversely polarized electrons from unpolarized radiation.[5] In practice not too much use has been made of the photoeffect for these purposes, due both to experimental problems and to the lack of precise knowledge of the expected correlations.[7]

We have recently completed a numerical calculation of K-shell photoelectric cross sections, covering the range of elements from $Z=13$ to $Z=92$ and the range of energies from 200 keV to 2 MeV, neglecting electron screening. Our numerical methods and the resulting differential and total cross sections have been reported elsewhere[8]; the purpose of this paper is to present and discuss the results which we have obtained for all the possible polarization correlations between incident photon and ejected electron. For a qualitative understanding of the nature of the correlations we give a set of figures showing the energy dependence of the correlations in a high-Z element; these are found in the separate sections in which we discuss each correlation. We have appended to this paper a representative set of tables for the element $Z=84$; a complete set of tables is also available.[9]

To define these correlations we need to introduce the usual polarization parameters. We describe photon polarization with the quantities

$$\xi_1=e_1^*e_1-e_2^*e_2\,,\quad \xi_2=e_1e_2^*+e_2e_1^*\,,\quad \xi_3=i(e_1e_2^*-e_2e_1^*)\,. \tag{1}$$

Here e_1 is the component of the photon polarization vector $\mathbf{e}$ which lies in the plane of the incoming photon $\mathbf{k}$ and outgoing electron $\mathbf{p}$, and e_2 is the component perpendicular to the plane (along $\mathbf{k}\times\mathbf{p}$), so that $(e_1,e_2,\mathbf{k})$ form a righthanded set. Photons linearly polarized parallel or perpendicular to the production plane are characterized by a nonvanishing ξ_1, circularly polarized photons by nonvanishing ξ_3. Similarly, we describe the polarization state of the ejected electron in terms of the direction $\boldsymbol{\zeta}$ of its spin in its rest system; ζ_3 is taken along the *electron* direction $\mathbf{p}$, ζ_1 in the scattering plane, and ζ_2 perpendicular to the scattering plane (along $\mathbf{k}\times\mathbf{p}$), so that $(\zeta_1,\zeta_2,\zeta_3)$ again form a right-handed set. Longitudinally polarized electrons are characterized by nonvanishing ζ_3. It should be noted that the coordinate systems used for the description of photon and electron polarizations are different.

In terms of these parameters the differential cross section for K-shell photoeffect, summed over both K-shell electrons, is of the form

$$\frac{d\sigma}{d\Omega}=\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpol.}}\left[\tfrac{1}{2}\sum_{i,j=0}\xi_i\zeta_jC_{ij}\right], \tag{2}$$

* This study was supported in part by the United States Air Force through Air Force Office of Scientific Research Grant AF-AFOSR-62-452, and in part by the United States Atomic Energy Commission.

† Address after August 1964: Department of Physics, University of Pittsburgh, Pittsburgh 13, Pennsylvania.

‡ Present address: Physics International, Berkeley, California.

[1] W. Heitler, *Quantum Theory of Radiation* (Oxford University Press, New York, 1954), 3rd ed.

[2] H. Olsen, Kgl. Norske Videnskab. Selskabs. Forh. **31**, No. 11 (1958).

[3] H. Banerjee, Nuovo Cimento **11**, 220 (1959).

[4] U. Fano, K. W. McVoy, and J. R. Albers, Phys. Rev. **116**, 1147 (1959); **116**, 1159 (1959).

[5] B. Nagel, Arkiv Fysik **18**, 1 (1960).

[6] R. H. Pratt, Phys. Rev. **123**, 1508 (1961).

[7] L. W. Fagg and S. S. Hanna, Rev. Mod. Phys. **31**, 711 (1959).

[8] R. H. Pratt, R. D. Levee, R. L. Pexton, and W. Aron, preceding paper, Phys. Rev. **134**, A898 (1964).

[9] These tables are available in limited supply from R. H. Pratt as Stanford University Institute of Theoretical Physics Report ITP-99 (unpublished). This report has also been deposited as Document No. 7829 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D. C. A copy may be secured by citing the Document number and by remitting $7.50 for photoprints, or $2.75 for 35 mm

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where $\xi_0=\zeta_0=C_{00}=1$, and $(d\sigma/d\Omega)_{\text{unpol.}}$ is the differential cross section from unpolarized photons, summed over final electron spins. The quantities C_{ij}, satisfying $|C_{ij}|\leq 1$, are the polarization correlations with which this paper is concerned.

It would appear from Eq. (2) that, not counting $C_{00}\equiv 1$, there are fifteen polarization correlations. But in fact invariance considerations permit only seven nonzero correlations. This follows most directly from the fact that the matrix element

$$H=-e(2\pi/k)^{1/2}\int d^3r\psi_{\text{fin}}{}^{\dagger}\boldsymbol{\alpha}\cdot\mathbf{e}e^{i\mathbf{k}\cdot\mathbf{r}}\psi_{\text{in}} \tag{3}$$

is invariant under the substitutions

$$\psi\rightarrow\sigma_2\psi\,,\quad (e_1,e_2)\rightarrow(-e_1,e_2)\,. \tag{4}$$

For we may absorb these substitutions into the polarization parameters by

$$(\xi_1,\xi_2,\xi_3)\rightarrow(\xi_1,-\xi_2,-\xi_3)\,,$$
$$(\zeta_1,\zeta_2,\zeta_3)\rightarrow(-\zeta_1,\zeta_2,-\zeta_3)\,, \tag{5}$$

and similarly for the bound-state polarization parameters which are summed out. If Eq. (2) is to be invariant under the substitutions (5), then we conclude that, not counting $C_{00}\equiv 1$, the only nonzero correlations are

$$C_{02},\,C_{10},\,C_{12},\,C_{21},\,C_{23},\,C_{31},\,C_{33}\,. \tag{6}$$

These seven correlations are indeed all nonzero, and they will be discussed in the subsequent sections. C_{02} produces transversely polarized electrons from unpolarized photons, C_{10} is the well-known correlation which analyzes linearly polarized photons, and C_{33} converts circularly polarized photons to longitudinally polarized electrons.

We have written down in Eqs. (2.25)–(2.27) of Ref. 8 expressions for the C_{ij} in terms of reduced partial-wave matrix elements R_κ and continuum wave-function phase shifts δ_κ; it is from these equations that the results presented in this paper were calculated. Some information which can be deduced from these equations concerning the behavior of the correlations at small angles should be mentioned. First of all, since the angular distribution $(d\sigma/d\Omega)_{\text{unpol.}}$ does not vanish for forward (and backward) angles (unlike the Sauter distribution,[1,10] valid only to lowest order in $a\equiv Z\alpha$), one sees that all correlations except C_{33} vanish in the forward and backward directions, and $C_{33}(0°)=+1$, $C_{33}(180°)=-1$. Further, for small angles the correlations C_{02}, C_{12}, C_{21}, and C_{31} go as θ, while C_{10} and C_{23} go as θ^2. And finally, for small angles one has the identity $C_{12}=-C_{21}$. However, one can see from the figures and tables that small angles, at these energies, means much less than 5°, and indeed all the correlations undergo very rapid and violent changes over this region of very small angles, in contrast with their comparatively slow variations through the main angular regions. Indeed, in the figures, we have not attempted to draw in the variation of the correlations between 0 and 5°.

Before proceeding to the specific correlations we may also discuss their behavior at low and at high energies and as a function of Z. In the nonrelativistic theory, $\boldsymbol{\alpha}\cdot\mathbf{e}$ in Eq. (3) is replaced by $\mathbf{p}\cdot\mathbf{e}$, and all dependence on electron spin disappears. This leads to a differential cross section proportional to $(1+\xi_1)$, so that $C_{10}=1$ for all angles and all other correlations vanish. Of course even at threshold there are relativistic corrections of relative order a^2 [which for example cause $C_{10}(0°)=0$], and away from threshold there are relativistic corrections of relative $O(1)$, as given by Sauter,[1,10] which in fact for the larger angles cause the preferred direction of emission to be perpendicular to rather than in the plane of $\mathbf{k}$ and $\mathbf{e}$. In the relativistic theory correlations with the electron spin can occur. In addition to C_{10} the correlations C_{31} and C_{33} are also $O(1)$, while the remaining correlations are $O(a)$ and so mainly important only in heavy elements. At very high energies one can show[6,11] that for angles of significant electron emission the only nonvanishing correlations are C_{12}, C_{21}, and C_{33}, and further $C_{33}=+1$ and $C_{12}=-C_{21}$.

II. THE CORRELATION C_{02}

The correlation C_{02} permits the production of transversely polarized electrons from unpolarized photons; the resulting electron spin is perpendicular to the production plane. The correlation vanishes both in the nonrelativistic and extreme relativistic limits and is $O(a)$. Figure 1 shows C_{02} for $Z=92$ at some representative energies. Like the more familiar correlation C_{10}, C_{02} exhibits a "crossover" feature: For energies above some minimum energy the spin direction $\mathbf{k}\times\mathbf{p}$ is favored at smaller angles and $-\mathbf{k}\times\mathbf{p}$ at larger angles. Previous work on this correlation is due to Kolbenstvedt and Olsen[12] and Nagel,[5] who obtained the analytic expression for the correlation to lowest nonvanishing order in a, and show curves which are qualitatively similar to ours. [These curves fail near the forward direction where the angular distribution itself is $O(a^2)$ rather than $O(1)$; the same comment will apply to all the analytical calculations subsequently mentioned.] Numerical results at threshold have been reported by Nagel and Olsson.[13] In the report on their numerical calculation of K-shell cross sections Hultberg, Nagel, and Olsson[14] show one representative curve for C_{02}

---

microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

[10] F. Sauter, Ann. Physik **11**, 454 (1931).

[11] B. Nagel, Arkiv Fysik **24**, 151 (1963).

[12] H. Kolbenstvedt and H. Olsen, Proceedings of the Physics Seminar in Trondheim, 1960 (unpublished).

[13] B. Nagel and P. Olsson, Arkiv Fysik **18**, 29 (1960).

[14] S. Hultberg, B. Nagel, and P. Olsson, Arkiv Fysik **20**, 555 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

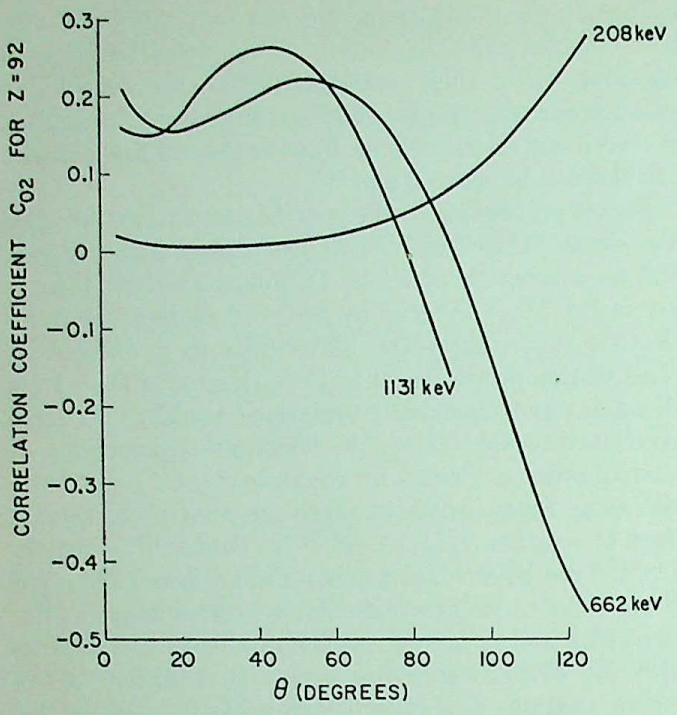


FIG. 1. Correlation C_{02} for the production of transversely polarized electrons from unpolarized photons, as a function of angle for $Z=92$ at several energies.

(P_1 in their notation), and in unpublished work[15] they have given curves for $Z=82$ and $Z=92$ over a range of energies up to 662 keV, which are in good agreement with the corresponding results of our tables. As has been noted,[5,12] the electron polarization which C_{02} implies is analogous to that produced in relativistic Coulomb scattering with unpolarized electrons, and can be analyzed in the same way; this correlation has the advantage[12] of not vanishing as rapidly at high energies.

III. THE CORRELATION C_{10}

The correlation C_{10}, as already noted, reflects the fact that the photoeffect can serve as an analyzer for photons linearly polarized parallel or perpendicular to the production plane. In the nonrelativistic theory[1] $C_{10}=+1$, which says no electrons are emitted perpendicular to the polarization direction of the photons. Of course, even at threshold there are relativistic corrections of $O(a^2)$, which, for example, cause $C_{10}(0°)=0$. The calculation of C_{10} to lowest order in a in the relativistic theory was made by Sauter,[1,10] and like C_{02} it exhibits a crossover feature: Above a minimum energy, emission in the plane of $\mathbf{k}$ and $\mathbf{e}$ (as predicted by the nonrelativistic theory) is favored for smaller angles of emission, whereas emission *perpendicular* to the plane is favored at larger angles. As discussed by Olsen,[2] in lowest order in a the minimum electron energy for the crossover feature is $\epsilon=\frac{5}{3}$ (in units $\hbar=c=m_e=1$), i.e., ≈ 340 keV for the photon, and at higher energies the crossover angle θ_c is determined from

$$k(k+1)(1-\beta\cos\theta_c)=2, \qquad (7)$$

where k is the photon energy in the same units. Olsen also noted that in the same approximation, at the crossover angle θ_c such that $C_{10}(\theta_c)=0$, electrons from *circularly* polarized photons are *completely* polarized along the *photon* direction. We shall return to this feature in our discussion of C_{31} and C_{33}. We should also note that the analytical corrections to C_{10} of order a were obtained by Gavrila[16] and Nagel.[5]

Results of our numerical calculation of C_{10} are given in the tables appended to this paper and in Fig. 2. We also give in Table I some information on the crossover angle θ_c as a function of k and Z. At low energies, the correlation is quite close to its nonrelativistic value except at large angles (and extremely close to the forward direction). As the energy increases the correlation appears to decrease monotonically for the angles of significant electron emission, and it will vanish in the high-energy limit. However, C_{10} can still be quite sizeable at 2 MeV. Table I shows that in heavy elements the crossover angle θ_c begins to vary significantly from the lowest order prediction Eq. (7); the crossover occurs significantly *later*.

Among the seven photoeffect polarization correlations, only C_{10} appears to have so far been the subject

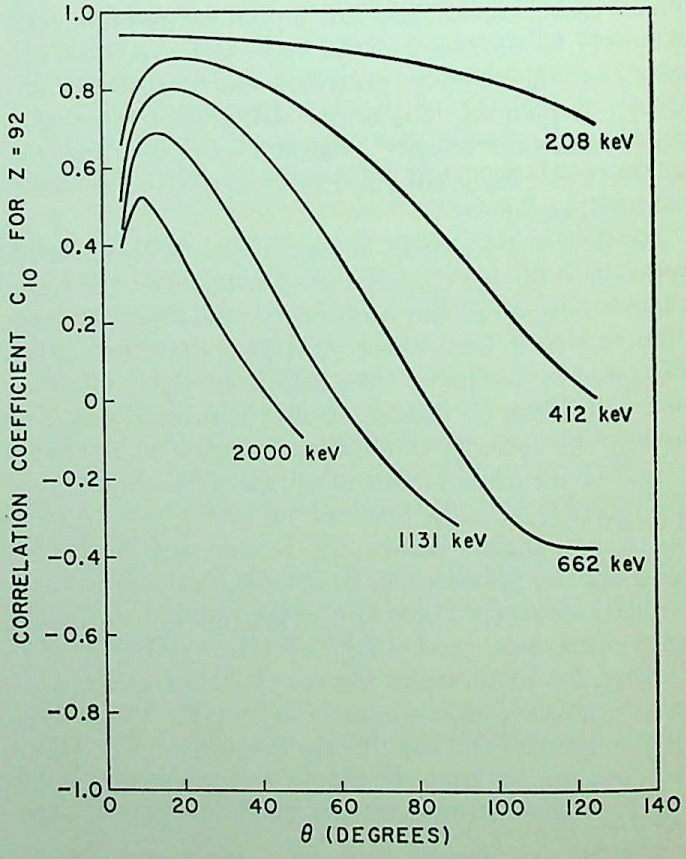


FIG. 2. Correlation C_{10} for the analysis of linearly polarized photons, as a function of angle for $Z=92$ at several energies.

[15] S. Hultberg, B. Nagel, and P. Olsson (unpublished). (Dr. Nagel now informs me that they have made calculations for all seven polarization correlations.)

[16] M. Gavrila, Phys. Rev. 113, 514 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of experimental investigations. In the energy range of interest there are experiments by Hereford and coworkers[17] and by Brini *et al.*[18] In both cases the objective was to test the predicted crossover feature of Sauter's relativistic theory; the first group appeared to see such a feature while the second group did not. Now that a more precise theory is available, and especially in view of the shift of the predicted θ_c as shown in Table I, further experiments on this question would seem desirable.

IV. THE CORRELATION C_{12}

The correlation C_{12} alters the production of transversely polarized electrons from that given by C_{02} in the case that the initial photons are linearly polarized parallel or perpendicular to the production plane. To determine C_{12} will require also determining C_{02} and C_{10}, or at least designing a set of experiments which subtract out their effect; C_{12} is the only one of the seven correlations which cannot in this sense be determined independently. Like C_{02}, the correlation C_{12} vanishes in the nonrelativistic limit and is $O(a)$; unlike C_{02} it remains finite in the extreme relativistic limit. Connections between C_{12} and C_{21} have already been mentioned: $C_{12}=-C_{21}$ for very small angles and also for all angles of significant electron emission (which are also small angles, but not as small) in the extreme relativistic limit.[6,11] There does not appear to be any discussion of the general analytic form of C_{12}, although an expression to lowest nonvanishing order in a could be extracted from Nagel's work.[5] An expression for the lowest order result valid in the extreme relativistic limit was obtained by Banerjee.[3] The complete analytical expression for the relativistic limit has now been given by Nagel,[11] and he has evaluated it numerically for three choices of Z. Results of our calculation of C_{12} are detailed in the Appendix and in Fig. 3. Since the correlation is $O(a)$ it is most important in heavy elements, but it can become quite sizeable for intermediate Z at back angles (and Nagel's work[11] indicates that in the extreme relativistic case it is greater for intermediate than for large Z). C_{12} is always positive in the regions considered here, but

TABLE I. Crossover angle θ_c for which $C_{10}(\theta_c)=0$ as a function of charge Z and photon energy k(MeV).

k \ Z	13	26	50	84
0.354	≈134°	≈132°	≈133°	≈157°
0.662	≈74°	≈68°	70°	80°
1.131	42°	41°	44°	56°
1.5	≈35°	≈35°	36°	47°
2.0	≈25°	≈24°	31°	40°

[17] F. L. Herford, Phys. Rev. **81**, 482 (1951); **81**, 627 (1951); F. L. Hereford and J. P. Keuper, *ibid.* **90**, 1043 (1953); W. H. McMaster and F. L. Hereford, *ibid.* **95**, 723 (1954).

[18] D. Brini, L. Pelli, O. Rimondi, and P. Veronesi, Nuovo Cimento **6**, 98 (1957).

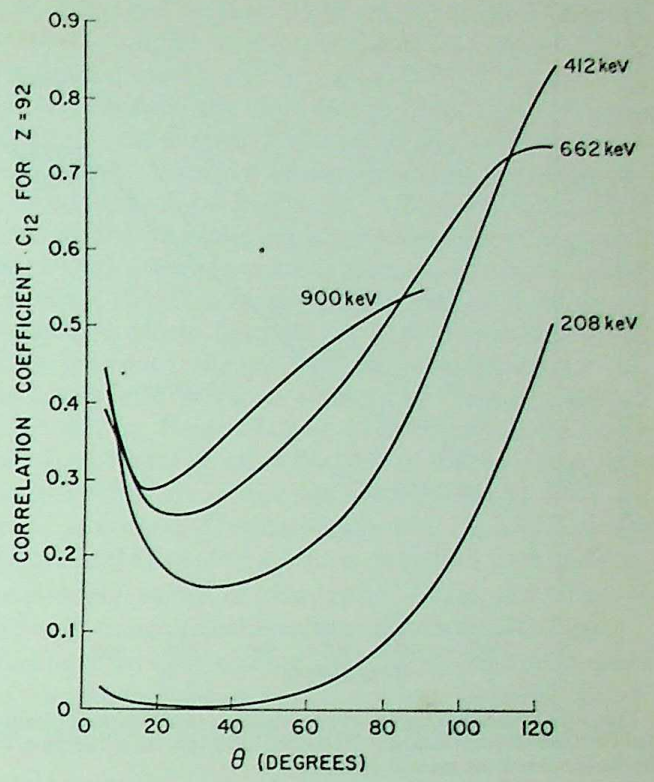


FIG. 3. Correlation C_{12} for the production of transversely polarized electrons from linearly polarized photons, as a function of angle for $Z=92$ at several energies.

Nagel's work[11] indicates that for very heavy Z and high energies there will be a crossover. Finally, for heavy Z and intermediate energies C_{12} is large, especially at the more backward angles, and so produces major effects in the production of transversely polarized electrons from linearly polarized photons.

V. THE CORRELATION C_{21}

The correlation C_{21}, like C_{12}, produces transversely polarized electrons from linearly polarized photons. However this correlation, also $O(a)$, is between electrons polarized *in* the production plane and photons polarized at 45° to the production plane. As we have said before, $C_{12}=-C_{21}$ for very small angles and also in the extreme relativistic limit. The remarks on previous work also apply to both processes: from Nagel's work[5] one can extract an analytic expression for the lowest nonvanishing order in a and from his more recent work[11] one of course has the extreme relativistic limit ($C_{21}=-C_{12}$, note there is a difference in sign and in his choice of coordinate system). Results for C_{21} are in the Appendix and in Fig. 4. Remarks on magnitudes apply to both correlations and need not be repeated here. However, the correlation C_{21} exhibits, even at the lowest energies, a crossover feature; the crossover angle increases with energy, unlike C_{02} or C_{10}. Since C_{21} is large at intermediate energies for back angles the photoeffect provides a fairly efficient mechanism for the transfer of these polarizations.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

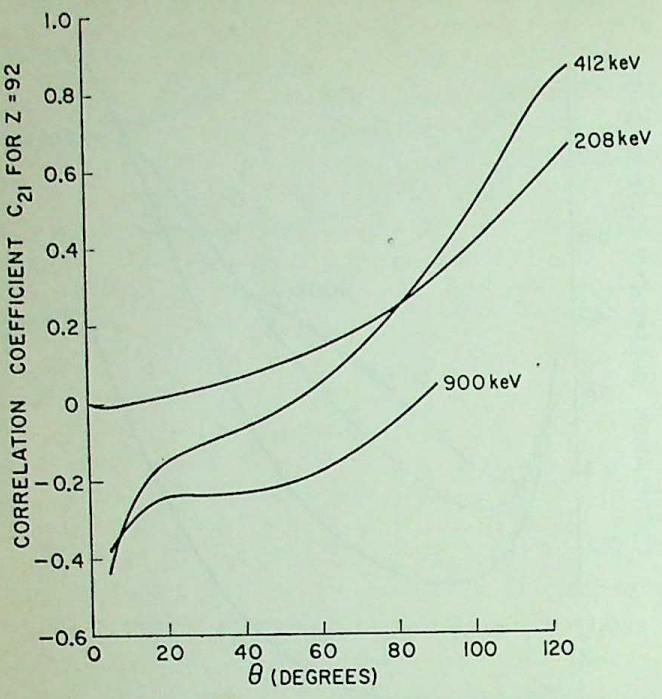


FIG. 4. Correlation C_{21} for the production of transversely polarized electrons from linearly polarized photons, as a function of angle for $Z=92$ at several energies.

VI. THE CORRELATION C_{23}

The correlation C_{23}, which is $O(a)$, produces longitudinally polarized electrons from linearly polarized photons. The photon polarization should again be at 45° to the production plane. There does not appear to have ever been any analytical discussion of this correlation, although the result to lowest order in a can again be extracted from Nagel's work.[5] Results for C_{23} are given in the Appendix and in Fig. 5. In the range considered here the correlation is always negative. And although C_{23} is $O(a)$ and vanishes in the nonrelativistic and extreme relativistic limits, for intermediate and large Z at intermediate energies it becomes quite sizeable at large angles and so provides a fairly efficient mechanism for the transfer of these polarizations.

VII. THE CORRELATION C_{31}

The correlation C_{31} produces electrons transversely polarized in the production plane from circularly polarized photons. The correlation vanishes in both the nonrelativistic and extreme relativistic limits, but it is $O(1)$ and very large for intermediate energies and angles. Figure 6 shows C_{31} for $Z=92$ at some representative energies; further results are given in the Appendix. At very forward angles the correlation is positive, but it soon crosses over and remains negative through most of the angular range, with a broad minimum at intermediate angles. (The curves suggest that there could actually be an energy, near 600 keV, and angle, near 80°, for which $C_{31}=-1$.) Analytical expressions for C_{31} to lowest order in a were obtained by Olsen[2] and Banerjee[3]; Nagel[5] also obtained terms of relative order a. Numerical results obtained by Hultberg, Nagel, and Olsson[14,15] (in their notation $C_{31}=-P_3$) are in good accord with the corresponding values of our tables. Numerical results at threshold were given by Nagel and Olsson.[13]

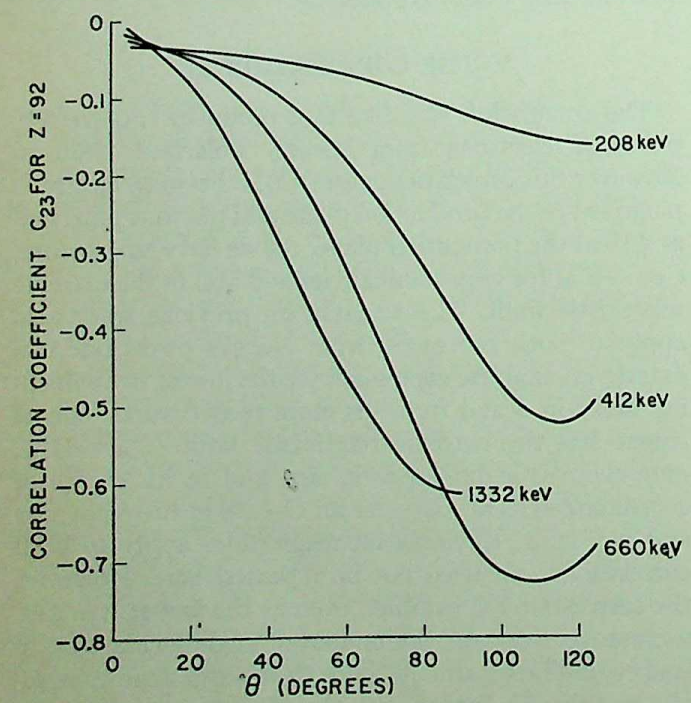


FIG. 5. Correlation C_{23} for the production of longitudinally polarized electrons from linearly polarized photons, as a function of angle for $Z=92$ at several energies.

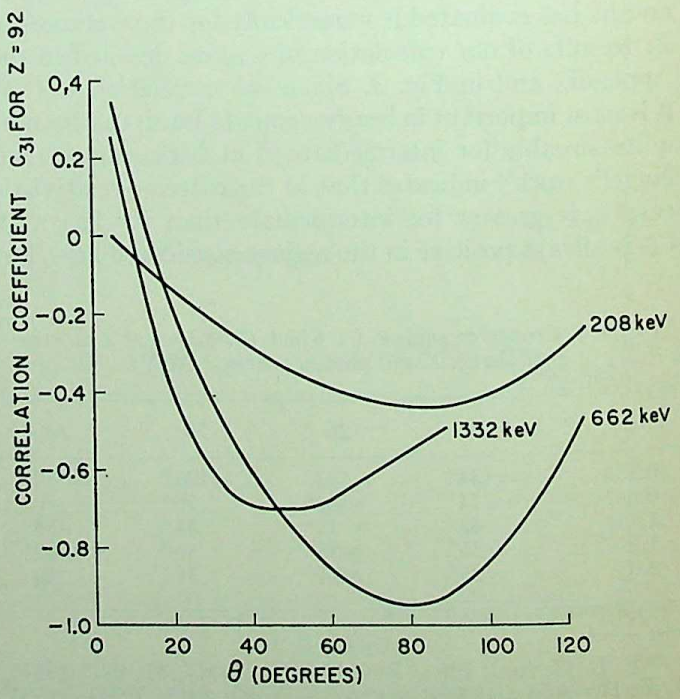


FIG. 6. Correlation C_{31} for the production of transversely polarized electrons from circularly polarized photons, as a function of angle for $Z=92$ at several energies.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. $D(\theta_c)$ as defined by Eq. (8) for $Z=84$ as a function of photon energy k (in MeV).

k	0.662	1.131	1.5	2.0
$D(\theta_c)$	0.963	0.972	0.965	0.961

It was noted by Olsen[2,4] that, to lowest order in a, at the crossover angle θ_c for which $C_{10}(\theta_c)=0$, electrons from circularly polarized photons are completely polarized along the *photon* direction, i.e.,

$$D(\theta_c)\equiv C_{33}(\theta_c)\cos\theta_c - C_{31}(\theta_c)\sin\theta_c = 1. \quad (8)$$

For this to be true requires that, in the language of Eqs. (2.25)–(2.27) of Ref. 8,

$$|K_-|^2=|J_+|^2=0 \text{ when } \mathrm{Re}[J_-{}^*J_+ + K_-{}^*K_+]=0; \quad (9)$$

it does not seem that this can in general be true. We show, however, in Table II the calculation of $D(\theta_c)$ for $Z=84$ from our results; the deviations from Olsen's prediction are very small, and almost within the errors of the calculation.

VIII. THE CORRELATION C_{33}

The correlation C_{33} is a transmitter of helicity, producing longitudinally polarized electrons from circularly polarized photons. The only one of the seven correlations which does not vanish in forward and backward directions [$C_{33}(0°)=-C_{33}(180°)=1$], it vanishes in the nonrelativistic limit, but becomes 1 for all angles of significant electron emission in the extreme relativistic limit.[6] We show results for C_{33} in the tables appended to the paper and in Fig. 7. Even at intermediate energies the correlation is large; it appears to approach one at high energies somewhat more rapidly in the lighter elements. Since C_{33} ranges from $+1$ to -1 it exhibits a crossover, which, however, at high energies moves to very backward angles. Analytical expressions for C_{33} to lowest order in a were obtained by Olsen,[2] by Banerjee,[3] and by Fano, McVoy, and Albers[4]; Nagel[5] also obtained terms of relative order a. The extreme relativistic limit of C_{33} was discussed by Pratt.[6] Like C_{02} and C_{31}, numerical results obtained for C_{33} by Hultberg, Nagel, and Olsson[14,15] are in good accord with the corresponding values of our tables. Nagel and Olsson[13] have given numerical results at threshold. The connection between the crossover angle for C_{10} and production of completely polarized electrons from circularly polarized photons has been discussed in the previous section.

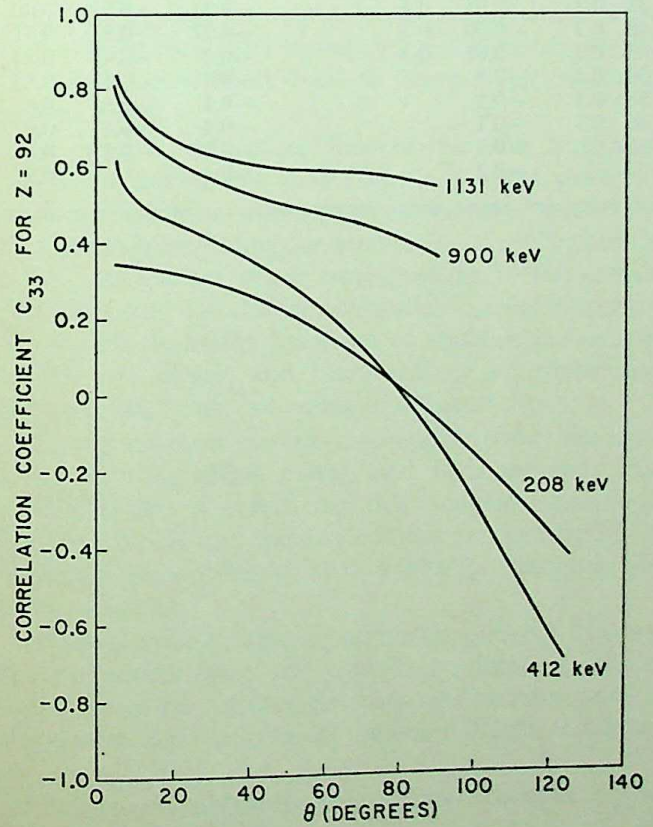


FIG. 7. Correlation C_{33} for the production of longitudinally polarized electrons from circularly polarized photons, as a function of angle for $Z=92$ at several energies.

ACKNOWLEDGMENTS

This work was suggested and encouraged by Dr. Harvey Hall, to whom all the authors are much indebted. Thanks are also due to Dr. Sölve Hultberg, for helpful discussions, correspondence, and prepublication results.

APPENDIX: TABLES III–VII

TABLE III. Polarization correlations for $Z=84$ at 354 keV as a function of angle.

θ (deg)	C_{02}	C_{10}	C_{12}	C_{21}	C_{23}	C_{31}	C_{33}
0	0.000	0.000	0.000	0.000	0.000	0.000	1.000
5	0.26	0.77	0.4	−0.4	−0.0	0.3	0.5
10	0.16	0.88	0.24	−0.2	−0.0	0.0	0.4
15	0.12	0.90	0.18	−0.16	−0.0	−0.0	0.4
20	0.10	0.90	0.15	−0.12	−0.0	−0.15	0.4
25	0.10	0.90	0.13	−0.09	−0.04	−0.22	0.38
30	0.09	0.88	0.12	−0.07	−0.05	−0.29	0.36
35	0.10	0.87	0.12	−0.05	−0.06	−0.36	0.34
40	0.10	0.85	0.12	−0.03	−0.07	−0.42	0.31
45	0.10	0.83	0.13	−0.01	−0.08	−0.48	0.28
50	0.10	0.81	0.13	0.01	−0.10	−0.53	0.25
55	0.11	0.78	0.14	0.03	−0.12	−0.58	0.21
60	0.11	0.75	0.16	0.06	−0.14	−0.63	0.18
65	0.11	0.72	0.18	0.09	−0.16	−0.67	0.13
70	0.12	0.68	0.20	0.13	−0.18	−0.71	0.08
75	0.12	0.65	0.22	0.18	−0.21	−0.74	0.03
80	0.12	0.61	0.26	0.22	−0.24	−0.76	−0.02
85	0.12	0.56	0.30	0.28	−0.27	−0.78	−0.08
90	0.12	0.51	0.35	0.34	−0.30	−0.78	−0.15
95	0.12	0.46	0.40	0.41	−0.33	−0.78	−0.22
100	0.12	0.41	0.46	0.48	−0.35	−0.76	−0.30
105	0.12	0.36	0.53	0.56	−0.38	−0.73	−0.38
110	0.11	0.31	0.60	0.64	−0.40	−0.69	−0.46
115	0.11	0.26	0.68	0.72	−0.41	−0.63	−0.54
120	0.10	0.21	0.75	0.79	−0.41	−0.56	−0.62
125	0.10	0.16	0.81	0.86	−0.41	−0.49	−0.70

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IV. Polarization correlations for $Z=84$ at 662 keV as a function of angle.

θ	C_{02}	C_{10}	C_{12}	C_{21}	C_{23}	C_{31}	C_{33}
0	0.000	0.000	0.000	0.000	0.000	0.000	1.000
5	0.218	0.57	0.40	−0.38	−0.04	0.294	0.74
10	0.177	0.78	0.30	−0.27	−0.04	0.058	0.62
15	0.161	0.81	0.264	−0.231	−0.03	−0.102	0.56
20	0.151	0.80	0.252	−0.221	−0.05	−0.222	0.54
25	0.144	0.77	0.241	−0.198	−0.07	−0.34	0.53
30	0.159	0.73	0.236	−0.157	−0.09	−0.46	0.50
35	0.190	0.69	0.250	−0.14	−0.10	−0.56	0.46
40	0.195	0.63	0.276	−0.16	−0.13	−0.64	0.44
45	0.174	0.56	0.292	−0.16	−0.18	−0.71	0.42
50	0.184	0.49	0.298	−0.1	−0.22	−0.78	0.38
55	0.225	0.42	0.316	−0.1	−0.24	−0.82	0.36
60	0.21	0.336	0.352	−0.1	−0.30	−0.87	0.34
65	0.14	0.233	0.386	−0.1	−0.36	−0.92	0.30
70	0.13	0.160	0.410	−0.1	−0.41	−0.94	0.26
75	0.16	0.109	0.42	0.1	−0.45	−0.93	0.24
80	0.12	0.001	0.46	0.1	−0.51	−0.94	0.22
85	−0.01	−0.140	0.52	0.1	−0.56	−0.95	;.15
90	−0.06	−0.196	0.54	0.2	−0.60	−0.92	0.08
95	−0.05	−0.18	0.55	0.3	−0.65	−0.86	0.07
100	−0.12	−0.28	0.59	0.3	−0.69	−0.82	0.04
105	−0.27	−0.4	0.66	0.4	−0.7	−0.76	−0.08
110	−0.34	−0.4	0.69	0.5	−0.7	−0.69	−0.19
115	−0.4	−0.4	0.68	0.5	−0.7	−0.63	−0.2
120	−0.4	−0.4	0.68	0.5	−0.7	−0.54	−0.25
125	−0.5	−0.4	0.7	0.6	−0.6	−0.43	−0.4
(deg)							

TABLE V. Polarization correlations for $Z=84$ at 1131 keV as a function of angle.

θ (deg)	C_{02}	C_{10}	C_{12}	C_{21}	C_{23}	C_{31}	C_{33}
0	0.000	0.000	0.000	0.000	0.000	0.000	1.000
5	0.18	0.48	0.42	−0.41	−0.019	0.263	0.82
10	0.15	0.67	0.34	−0.33	−0.035	0.007	0.79
15	0.15	0.68	0.30	−0.27	−0.051	−0.210	0.70
20	0.17	0.63	0.30	−0.25	−0.072	−0.379	0.67
25	0.19	0.56	0.32	−0.26	−0.098	−0.507	0.65
30	0.21	0.48	0.35	−0.27	−0.136	−0.602	0.63
35	0.23	0.38	0.37	−0.27	−0.18	−0.675	0.62
40	0.23	0.28	0.38	−0.24	−0.23	−0.73	0.61
45	0.22	0.19	0.38	−0.22	−0.27	−0.75	0.61
50	0.22	0.10	0.40	−0.22	−0.32	−0.76	0.61
55	0.20	0.01	0.4	−0.22	−0.39	−0.76	0.61
60	0.16	−0.06	0.4	−0.18	−0.43	−0.75	0.60
65	0.11	−0.13	0.4	−0.13	−0.45	−0.74	0.61
70	0.07	−0.19	0.4	−0.1	−0.49	−0.70	0.61
75	0.01	−0.24	0.4	−0.1	−0.55	−0.66	0.61
80	−0.06	−0.28	0.4	−0.1	−0.58	−0.64	0.59
85	−0.13	−0.32	0.4	0.0	−0.58	−0.62	0.57
90	−0.19	−0.36	0.4		−0.59	−0.57	0.56

TABLE VI. Polarization correlations for $Z=84$ at 1500 keV as a function of angle.

θ (deg)	C_{02}	C_{10}	C_{12}	C_{21}	C_{23}	C_{31}	C_{33}
0	0.000	0.000	0.000	0.000	0.000	0.000	1.000
5	0.15	0.45	0.43	−0.43	−0.01	0.22	0.85
10	0.14	0.61	0.35	−0.34	−0.03	−0.05	0.79
15	0.14	0.58	0.32	−0.29	−0.06	−0.29	0.76
20	0.18	0.49	0.32	−0.26	−0.09	−0.46	0.73
25	0.21	0.40	0.35	−0.28	−0.12	−0.56	0.72
30	0.22	0.30	0.40	−0.33	−0.16	−0.61	0.72
35	0.21	0.20	0.42	−0.32	−0.23	−0.65	0.72
40	0.22	0.11	0.38	−0.24	−0.28	−0.67	0.71
45	0.23	0.03	0.4	−0.2	−0.29	−0.64	0.73
50	0.19	−0.05	0.4		−0.33	−0.62	0.74
55	0.1	−0.1	0.4		−0.41	−0.62	0.73
60	0.1	−0.1	0.4		−0.4	−0.59	0.73
65		−0.2	0.3		−0.4	−0.53	0.76
70		−0.2	0.3		−0.4	−0.49	0.76

TABLE VII. Polarization correlations for $Z=84$ at 2000 keV as a function of angle.

θ (deg)	C_{02}	C_{10}	C_{12}	C_{21}	C_{23}	C_{31}	C_{33}
0	0.000	0.000	0.000	0.000	0.000	0.000	1.000
5	0.14	0.42	0.46	−0.46	−0.01	0.17	0.88
10	0.12	0.53	0.37	−0.36	−0.04	−0.11	0.83
15	0.13	0.45	0.32	−0.29	−0.08	−0.36	0.82
20	0.18	0.32	0.29	−0.23	−0.11	−0.51	0.80
25	0.22	0.23	0.35	−0.3	−0.13	−0.54	0.80
30	0.20	0.14	0.46	−0.4	−0.19	−0.55	0.80
35	0.2	0.05	0.4		−0.27	−0.57	0.80
40	0.2	−0.00	0.3		−0.27	−0.53	0.81
45	0.2	−0.07	0.3		−0.27	−0.47	0.84
50	0.1	−0.1			−0.3	−0.47	0.83
55	0.1	−0.1			−0.4	−0.5	0.8
60	0.1	−0.1			−0.4	−0.4	0.8
65		−0.2			−0.4	−0.5	0.8
70		−0.2			−0.4		0.8

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Effect of Correlation on the Ferromagnetism of Transition Metals

MARTIN C. GUTZWILLER

*IBM Zurich Research Laboratory, Ruschlikon, Zurich, Switzerland**

(Received 17 October 1963; revised manuscript received 13 December 1963)

The wave function for the electrons is investigated when a set of narrow bands (valence states) has its energies within a wide band (conduction states). The valence states are linear combinations of localized states which are attached to each lattice site. The intra-atomic Coulomb and exchange integrals for the localized states are much larger than the bandwidths of the valence states. Some of the narrow bands are neither completely empty nor completely filled. The wave function is therefore expected to be correlated, because it is disadvantageous for the electrons to crowd into the same lattice site, or take up some configuration contrary to Hund's rule. This correlation is important in transition metals, where it is considered to be the cause of ferromagnetism. The correlated wave function is obtained by applying to the uncorrelated antisymmetrized product of Bloch functions an operator which provides each configuration of localized valence states with an appropriate amplitude and phase factor. The procedure is worked out in detail for the case of few particles (electrons or holes) in the narrow bands with the help of a diagram analysis. The localized orbits of different lattice sites do not have to be orthogonal to on another, and the computational rules are actually simplified thereby. The example of a twofold degenerate band such as the upper part of the $3d$ band in Ni is treated, and the conditions for the occurrence of ferromagnetism are stated in the case of few $3d$ holes per lattice site.

INTRODUCTION

THE present work studies the correlation of electrons in a situation which is considered as typical for the $3d$ electrons in a transition metal. The characteristic features of the valence and the conduction electrons in the transition metals are not obtained from first principles, but they are assumed according to a number of experimental facts and their intuitive interpretation in order to yield a model Hamiltonian. The aim is therefore to find an approximate ground-state wave function for the model Hamiltonian, and to investigate the mechanisms which tend to favor a ferromagnetic ground state.

In accord with other theories for the transition metals[1] it is assumed that there is a fairly clear cut distinction between three types of electronic states: the core states representing the atomic $1s$, $2s$, $2p$, $3s$, and $3p$ orbits, the valence states corresponding to the atomic $3d$ orbits, and the conduction states corresponding to the atomic $4s$ orbits. Each set of states is orthogonal to the two others, and there will be no correlation between two orbits belonging to different sets.

This is the most obvious assumption about the electrons in a transition metal, and it is certainly not definitive. But it is felt that this complete separation between the $3d$ and the $4s$ electrons can be lifted in a later, refined model without basically changing the present results.

There is some striking experimental evidence in favor of the independence and completely different orbital character of the valence electrons on one hand and of the conduction electrons on the other. Neutron diffraction in iron[2] has shown that the magnetization density in the lattice is mostly concentrated very tightly around the lattice sites, whereas a small fraction of the total magnetization is distributed uniformly throughout the lattice. The simple linear dependence of magnetization on the alloy composition, e.g. in the Ni-Cu system,[3] seems to indicate that the Bloch states which carry most of the magnetization are being filled quite independently from the presence of conduction electrons.

The conduction electrons are formally included in the present investigation. Unlike the valence states which are defined in terms of tightly localized orbits, the conduction states may best be thought of as plane waves which were orthogonalized to the core and valence states. The lowest conduction state is assumed to lie well below all the valence states, and the highest conduction state well above. The noncrossing rule makes it inevitable to introduce bands which are conduction like in part of their Brillouin zone and valence like in the remaining part. In the transition region the Bloch states are a mixture of the two kinds, but all these transition regions together will be assumed to occupy a volume in reciprocal space which is small compared to the volume of one Brillouin zone. With this hypothesis all the actual computations will bear out the kinematic and dynamic independence of the conduction states from the valence states. The presence of the conduction electrons in the crystal manifests itself then mainly through the seemingly arbitrary number of valence electrons even in a "simple" metal, such as pure Fe, Co, and Ni. Other effects of the conduction electrons are harder to evaluate and will be either implicitly contained or completely neglected. The screening of the

* Present address: IBM Watson Laboratory, Columbia University, New York, New York.

[1] Cf. in particular the review article of J. H. van Vleck, Rev. Mod. Phys. 25, 220 (1953) which gives a qualitative description of the problems discussed in the present work.

[2] C. G. Shull and Y. Yamada, *Proceedings of the International Conference on Magnetism and Crystallography*, Vol. III, 1 (1961); J. Phys. Soc. Japan 17, Suppl. B (1962).

[3] H. C. van Elst, B. Lubach, and G. J. van den Berg, Physica 28, 1297 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

nuclear charge by the conduction electrons is implicitly taken into account through the effective crystal potential, but the coupling of the spins of $3d$ electrons through the conduction electrons is not mentioned at all since its treatment belongs to a more refined model.

The $3d$ electrons are characterized by the localized (at each lattice site) states which they occupy. No precise description of these atomic states is needed besides their transformation properties under the symmetry operations of the lattice, and the crucial assumption that two such localized orbits at different lattice sites have a small overlap integral. The smallness of this overlap is not a consequence of a clever choice, such as in a Wannier function, but a genuine expression for the tightness of the $3d$ orbits around their lattice site. Therefore, the matrix elements for a transition of a localized $3d$ electron to a neighboring lattice site due to kinetic energy and the crystal potential, are always smaller (by a factor of the order of the overlap integral) than the matrix elements for a transition within the same lattice site. The transitions between lattice sites due to the Coulomb repulsion or exchange between electrons are smaller by the square of the overlap integral than the transition within one lattice site. The only transitions between lattice sites which are admitted by the model Hamiltonian are therefore the ones due to the kinetic energy and the crystal potential. Those terms determine the band structure of the $3d$ electrons. The bandwidth is then smaller than the intra-atomic Coulomb and exchange integrals.

There are three terms in the model Hamiltonian: (i) kinetic energy and crystal potential terms, (ii) Coulomb repulsion terms at the same lattice site, (iii) exchange terms at the same lattice site. Terms involving the integral over four orbits at the same lattice site, but not of the ordinary Coulomb repulsion or the ordinary exchange type, are neglected. They are not necessarily small, but it is believed that the important ordering of atomic levels at one particular lattice site can already be understood on the basis of the Coulomb repulsion and exchange integrals. In particular Hund's rule for an atomic system follows already from such a reduced Hamiltonian.

Since the two particle terms (ii) and (iii) are larger than the one particle terms (i), the resulting problem is one of strong interaction. It is, however, complicated by the nonvanishing overlap between orbits on different lattice sites. The usual formulation in terms of creation and annihilation operators for the localized states has to be modified. The destruction operator for a given state is not exactly the Hermitian conjugate of its creation operator. Although this complication does not arise in the usual formulation, it does not really make the computations more awkward. It turns out that the various expectation values for the correlated wave function can be evaluated only in the limit of few particles (electrons or holes) compared to the number of lattice sites. The nonvanishing overlap integrals help to simplify the general rules in the diagrammatic analysis, although the number of diagrams is increased.

The present investigation amplifies considerably a short account published earlier.[4] The main idea is still the same. The uncorrelated antisymmetrized product of Bloch states is modified by applying operators which essentially project the uncorrelated wave function into a subspace of configurations with the appropriate correlations. The resulting correlated wave function is believed to be a good approximation to the ground state for all densities of electrons or holes. Practical results, however, i.e., the calculation of various expectation values, have been possible only in the limit of low-particle densities (with the exception of one statement concerning the occupation probability for valence states inside the Fermi surface in the case of a nondegenerate band, cf. Ref. 4). In the earlier account there was neither the possibility of nonvanishing overlap, nor of degenerate bands, nor of conduction states, nor of combining configurations according to Hund's rule. All these new features are brought out in the example of a twofold degenerate band, where ferromagnetism is favored by both terms, (ii) and (iii).

For the nonferromagnetic state the intra-atomic Coulomb-repulsion terms (ii) are more effective in scattering the particles into Bloch states of high energy, because two uncorrelated particles crowd more often into the same lattice site if their spins are antiparallel. Although this effect is probably not serious enough to induce ferromagnetism, it may come close to canceling the loss in band energy which particles of the same spin experience due to the exclusion principle. The dependence of this mechanism on the density of states was shown in the earlier account, and is again found in the example at the end of this report. But now, there is also the effect of the intra-atomic exchange which had been pointed out by Slater[5] and which is due to the terms (iii). Whereas one may hope to show some day that the terms (ii) can never induce ferromagnetism at all by themselves, it would still seem highly arbitrary to neglect them altogether in the discussion of ferromagnetism and consider only the terms (iii).

A brief account of the following sections will now be given. The first- and second-order density functions are defined in Sec. 1 together with the relevant overlap integrals. The creation and destruction operators are introduced in Sec. 2 and the important projection operators for the localized valence states are studied. The basic idea for the definition of a correlated wave function is explained in Sec. 3 by studying the correlation energy due to intra-atomic Coulomb repulsion in the simplest possible case of all Coulomb integrals being equal. The correlated wave function is defined in Sec. 4 quite generally by applying a suitable set of operators (projections operators for the various states of the whole

[4] Martin C. Gutzwiller, Phys. Rev. Letters **10**, 159 (1963).
[5] J. C. Slater, H. Statz, and G. F. Koster, Phys. Rev. **91**, 1323 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ion of each lattice site) to the uncorrelated antisymmetrized product of Bloch states. The same idea is realized in Sec. 5 in a slightly different manner, which will then help to expand all expectation values in powers of the particle density.

The computation of expectation values for the correlated wave function is outlined in Sec. 6; it leads to evaluating the expectation value of a product of creation and destruction operators for the uncorrelated wave function. The latter expectation values are represented in Sec. 7 by diagrams which contain vertices, solid lines for particle propagators, and dotted lines for overlap integrals. The usual breaking up of a diagram into closed linked diagrams is performed in Sec. 8; also it is shown how to obtain diagrams with external points from closed linked diagrams.

The first-order density function is studied in Sec. 9 in order, first to eliminate a number of simple diagrams which lead to a renormalization of the initial uncorrelated wave function, and second, to obtain all the terms up to second order in the particle density. The second-order density function is similarly treated in Sec. 10. The main correlation parameters occur at most quadratically and can therefore easily be determined by minimizing the expectation value for the total energy. This procedure is applied in Sec. 11 to the twofold-degenerate band characteristic of the upper end of the $3d$ band. The conditions for ferromagnetism are then stated. The density of states per unit interval of energy at the upper end of the valence band has to be larger than the average density in the limit of very strong localized Coulomb repulsion. The degeneracy of the valence band is important in the more realistic case where the Coulomb repulsion is not infinitely larger than the bandwidth so that two holes can occasionally crowd into the same lattice site.

The results of the last section are derived for a small density of valence holes and are based on the assumed correlated wave function. This wave function constitutes a guess which may not be particularly good for low densities, because the correlation may spread over many neighbors for low-particle densities as opposed to densities of about 1 per atom. Since the low-density case can be treated directly with the help of the ladder approximation, the present investigation may be viewed as an attempt to propose a wave function for all densities and to treat it in the limit where exact results can be obtained.

1. THE SINGLE-ELECTRON WAVE FUNCTIONS

The lattice is made up of L identical atoms in some elementary arrangement, say face-centered cubic. The lattice sites are labeled by the letters f, g, or h. To each site belongs a set of single-electron wave functions, which we shall divide into three categories: core states (corresponding to $1s$, $2s$, $2p$, $3s$, $3p$ levels); valence states (corresponding to $3d$ levels); and conduction states (corresponding to $4s$ levels).

The core states are assumed orthogonal to the valence and conduction states of all atomic sites. Also, the core states are uncorrelated, so that they contribute only through the first-order density function $\rho_1(x,y)$, namely

$$\rho_1(x,y)=N\int dx_2\cdots\int dx_N \times\Psi(x,x_2,\cdots,x_N)\Psi^*(y,x_2,\cdots,x_N) =\sum\chi(x)\chi^*(y)+\text{contribution from valence states and conduction states,} \quad (1)$$

where the first term on the right is from the core states.

In this formula the core states χ may be either localized or Bloch-like. The contribution of the valence and conduction states to ρ_1 is complicated because of correlation. The contribution of the core states to the second-order density function $\rho_2(\xi,x;\eta,y)$ is then already contained in

$$\rho_2(\xi,x;\eta,y)=\tfrac{1}{2}N(N-1)\int dx_3\cdots dx_N \times\Psi(\xi,x,x_3,\cdots,x_N)\Psi^*(\eta,y,x_3,\cdots,x_N) =\tfrac{1}{2}[\rho_1(\xi,\eta)\rho_1(x,y)-\rho_1(\xi,y)\rho_1(x,\eta)] +\text{properly correlated part arising from valence and conduction states.} \quad (2)$$

Since the effects of the core states are completely described by (1) and (2) together with their orthogonality to the valence and conduction states, it is quite feasible not to mention the core states any more in working out a correlated wave function for the valence and conduction states. Therefore, the first term in (1) will be mentioned again only when it comes to computing the total energy of the crystal. In keeping with this procedure we shall now call N the number of electrons in valence and conduction states only, and the total wave function Ψ will have only as many variables x as there are electrons in valence and conduction states. Formulas (1) and (2) are interpreted henceforth without the first term on the right-hand side of (1). The eventual inclusion of core states is easily performed.

The localized valence and conduction states are $\varphi_{\beta g}(x)$, where g is the label of the atomic site, β is the band index (including spin index), and x is the combination of spatial and spin variable. The band index is sometimes labeled α or γ. Bloch states are constructed by

$$\psi_{\kappa k}(x)=(1/L^{1/2})\sum_{\beta g}V_{\kappa k,\beta g}\varphi_{\beta g}(x). \quad (3)$$

k is a wave vector of the first Brillouin zone and κ is the index of the band. The coefficient $V_{\kappa k,\beta g}$ depends on the lattice site g through a factor $\exp(ikg)$, but the dependence on the indices κ and β may be quite complicated. In particular, $\psi_{\kappa k}$ may contain, for certain combinations

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of κ and k, contributions from both valence and conduction states. The exact values of $V_{\kappa k,\beta g}$ will result from some secular problem, exactly as the band structure is obtained in the tight-binding approximation. The variety of Bloch states $\psi_{\kappa k}$ depends directly on the initial choice of localized valence and conduction states $\varphi_{\beta g}$.

Formula (3) can be inverted by writing

$$\varphi_{\beta g}(x)=(1/L^{1/2})\sum_{\kappa k} U_{\beta g,\kappa k}\psi_{\kappa k}(x)\,, \tag{4}$$

where one has the relations

$$\begin{aligned}(1/L)\sum_{\kappa k} U_{\alpha f,\kappa k}V_{\kappa k,\gamma h}&=\delta_{\alpha\gamma}\delta_{fh}\,,\\ (1/L)\sum_{\beta g} V_{\lambda l,\beta g}U_{\beta g,\kappa k}&=\delta_{\lambda\kappa}\delta_{lk}\,.\end{aligned} \tag{5}$$

The coefficients $U_{\beta g,\kappa k}$ depend on g through the factor $\exp(-ikg)$.

The Bloch functions $\psi_{\kappa k}(x)$ will be assumed orthonormal to one another, i.e.,

$$\int \psi_{\lambda l}^{*}(x)\psi_{\kappa k}(x)dx=\delta_{\lambda\kappa}\delta_{lk}\,. \tag{6}$$

But the localized orbitals $\varphi_{\beta g}(x)$ will not be assumed orthonormalized, and this makes it necessary to introduce their overlap integrals

$$\begin{aligned}R(\alpha f,\gamma h)&=\int \varphi^{*}_{\alpha f}(x)\varphi_{\gamma h}(x)dx\\ &=(1/L)\sum_{\kappa k} U^{*}_{\alpha f,\kappa k}U_{\gamma h,\kappa k}\,.\end{aligned} \tag{7}$$

The only restriction which may be imposed for convenience is

$$R(\alpha f,\gamma f)=\delta_{\alpha\gamma}\,, \tag{8}$$

so that the localized orbits belonging to the same lattice site are orthonormalized. Also, the values of $R(\alpha f,\gamma h)$ differ from zero only when the sites of f and h are close to each other, and then they are small compared to 1 in all cases except when the indices α and γ refer both to conduction states.

Two remarks are in order at this point. First, it is convenient to introduce the nonorthogonality integrals R, even if they should be ultimately assumed to vanish. Indeed, the results of the diagrammatic analysis of Sec. 7 are more easily expressed if the values of $R(\alpha f,\gamma h)$ are not immediately simplified to $\delta_{\alpha\gamma}\delta_{fh}$. Second, it may be that certain simple types of correlations, special cases of the ones to be introduced in Sec. 4, do not necessitate the assumption of $R(\alpha f,\gamma h)$ being different from $\delta_{\alpha\gamma}\delta_{fh}$. Indeed the nonorthogonality of the valence states may be taken fully into account by a proper choice of the correlation parameters. An example is the Heitler-London theory of the hydrogen molecule. But no such equivalence between correlation parameters and nonorthogonality integrals has been found in the more general case to be discussed in the following sections for the Ni-Cu alloys.

2. SECOND QUANTIZED FORMALISM

The many-electron wave functions may be treated most conveniently in the formalism of second quantization. Therefore, we introduce creation operators $c^{\dagger}_{\kappa k}$ and annihilation operators $c_{\kappa k}$ corresponding to the Bloch functions $\psi_{\kappa k}(x)$ of (3). These operators satisfy the anticommutation relations

$$\begin{aligned}c^{\dagger}_{\kappa k}c_{\lambda l}+c_{\lambda l}c^{\dagger}_{\kappa k}&=\delta_{\kappa\lambda}\delta_{kl}\,,\\ c_{\kappa k}c_{\lambda l}+c_{\lambda l}c_{\kappa k}&=0\,,\\ c^{\dagger}_{\kappa k}c^{\dagger}_{\kappa k}+c^{\dagger}_{\kappa k}c^{\dagger}_{k}&=0\,.\end{aligned} \tag{9}$$

The localized states $\varphi_{\beta g}(x)$ are created by

$$a^{\dagger}_{\beta g}=\frac{1}{L^{1/2}}\sum_{\kappa k} U_{\beta g,\kappa k}c^{\dagger}_{\kappa k}\,, \tag{10}$$

and they are destroyed by the operators

$$b_{\beta g}=\frac{1}{L^{1/2}}\sum_{\kappa k} V_{\kappa k,\beta g}c_{\kappa k}\,. \tag{11}$$

It is important to notice that $b^{\dagger}_{\beta g}\neq a^{\dagger}_{\beta g}$ because the localized orbits are not orthonormalized. However, one has

$$\begin{aligned}a^{\dagger}_{\alpha f}b_{\gamma h}+b_{\gamma h}a^{\dagger}_{\alpha f}&=\delta_{\alpha\gamma}\delta_{fh}\,,\\ a^{\dagger}_{\alpha f}a^{\dagger}_{\gamma h}+a^{\dagger}_{\gamma h}a^{\dagger}_{\alpha f}&=0\,,\\ b_{\alpha f}b_{\gamma h}+b_{\gamma h}b_{\alpha f}&=0\,.\end{aligned} \tag{12}$$

These relations follow from (9) and (5). Other anticommutators between a's and b's can be expressed in terms of the overlap integral R of (7).

A one-particle operator $Q_1(x,y)$ with matrix elements

$$\langle\kappa k|Q_1|\lambda l\rangle=\int\psi^{*}_{\kappa k}(x)dx\int\psi_{\lambda l}(y)dyQ_1(x,y)\,, \tag{13}$$

is given in second quantized formalism by

$$Q_1=\sum_{\kappa k,\lambda l} c^{\dagger}_{\kappa k}\langle\kappa k|Q_1|\lambda l\rangle c_{\lambda l}\,. \tag{14}$$

In terms of the matrix elements

$$\langle\alpha f|Q_1|\gamma h\rangle=\int\varphi^{*}_{\alpha f}(x)dx\int\varphi_{\gamma h}(y)dyQ_1(x,y)\,, \tag{15}$$

the operator Q_1 becomes in the second quantized formalism

$$Q_1=\sum_{\alpha f,\gamma h} b^{\dagger}_{\alpha f}\langle\alpha f|Q_1|\gamma h\rangle b_{\gamma h}\,, \tag{16}$$

which is obtained directly from (13) and (14) by inserting (3) and (11).

In the same manner, a two-particle operator

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$Q_2(\xi,x;\eta,y)$ with matrix elements

$$\langle\alpha_1 f_1\alpha_2 f_2|Q_2|\gamma_1 h_1\gamma_2 h_2\rangle=\tfrac{1}{2}\int d\xi\int dx\int d\eta\int dy \times\begin{vmatrix}\varphi^*_{\alpha_1 f_1}(\xi) & \varphi^*_{\alpha_2 f_2}(\xi)\\ \varphi^*_{\alpha_1 f_1}(x) & \varphi^*_{\alpha_2 f_2}(x)\end{vmatrix} Q_2(\xi,x;\eta,y) \times\begin{vmatrix}\varphi_{\gamma_1 h_1}(\eta) & \varphi_{\gamma_2 h_2}(\eta)\\ \varphi_{\gamma_1 h_1}(y) & \varphi_{\gamma_2 h_2}(y)\end{vmatrix} \quad (17)$$

is given in second quantized form by the formula

$$Q_2=\sum_{(\alpha_1 f_1,\alpha_2 f_2)}\sum_{(\gamma_1 h_1,\gamma_2 h_2)} b^\dagger_{\alpha_2 f_2} b^\dagger_{\alpha_1 f_1} \times\langle\alpha_1 f_1\alpha_2 f_2|Q_2|\gamma_1 h_1\gamma_2 h_2\rangle b_{\gamma_1 h_1} b_{\gamma_2 h_2}. \quad (18)$$

The summation goes only over different pairs $(\alpha_1 f_1,\alpha_2 f_2)$ and $(\gamma_1 h_1,\gamma_2 h_2)$. The formulas (16) and (18) have been derived because in this work the assumption of orthonormality is not made for the localized states, so that there remains the question whether the a or the b operators are to be used in representing Q_1 and Q_2.

The operator $a^\dagger_{\beta g} b_{\beta g}=P_{\beta g}$ destroys the electron in the localized state $\varphi_{\beta g}(x)$ and then creates right back again. This operator acting on an arbitrary many electron state isolates that part which has an electron in the localized state $\varphi_{\beta g}(x)$. In view of

$$P^2_{\beta g}=a^\dagger_{\beta g} b_{\beta g} a^\dagger_{\beta g} b_{\beta g}=a^\dagger_{\beta g} b_{\beta g}=P_{\beta g}, \quad (19)$$

the operator $P_{\beta g}$ is indeed a projection operator. However, it is not a Hermitian operator, and has to be handled with some care. One way of examining it in more detail is to express $a^\dagger_{\beta g}$ in terms of $b^\dagger_{\beta g}$. With the help of (5), (10), and (11) one finds

$$a^\dagger_{\beta g}=\sum_{\alpha f} b^\dagger_{\alpha f} R(\alpha f,\beta g), \; =b^\dagger_{\beta g}+\sum_{\alpha f\neq g} b^\dagger_{\alpha f} R(\alpha f,\beta g). \quad (20)$$

The last step is a consequence of the orthonormalization (8) for the localized states belonging to the same lattice site. If the index β refers to a valence state, the second term on the right hand side of (20) is small to the extent that the overlaps between a localized valence state and all localized states on neighbors are small. The overlap between a valence and a conduction state is small because of the different angular symmetry, and the overlap between two valence states is small because of their tight localization. In this sense the operator $P_{\beta g}$ is approximately Hermitian for the localized valence states.

Since the projection operator $P_{\beta g}$ is not Hermitian we cannot associate a measurable quantity with it. As long as $P_{\beta g}$ is only used to construct some correlated wave function for the whole crystal, the lack of Hermiticity is not a drawback. But $P_{\beta g}$ cannot be used to answer the question as to the probability of finding the localized valence state $\varphi_{\beta g}$ occupied. Such detailed knowledge may be of no interest. However, one may find it useful to answer the more general question as to the average number of electrons in localized valence states. A Hermitian operator whose expectation value answers the last question can be constructed under certain additional assumptions which do not seem unduly restrictive. For instance, the conduction states may be defined as plane waves which were orthogonalized not only to the core states but also to the valence states. An operator can then be constructed whose expectation value gives the average number of electrons in conduction states, and therefore the difference between the total number of electrons and the number of electrons in conduction states gives the number of electrons in valence states. We shall use such a number in the future whenever this seems convenient.

3. ABOUT THE CORRELATION IN THE VALENCE STATES

With the help of the projection operator $P_{\beta g}$ the contribution of the localized valence state $\varphi_{\beta g}$ to the wave function Ψ of the whole crystal can be investigated. More generally, one can speak about the configurations Φ which contribute to Ψ. A configuration is a set of localized valence states, and it can be described by enumerating the index pairs βg which make up the configuration. To each configuration Φ belongs a projection operator which is just the product of $P_{\beta g}$ for the indices belonging to the configurations Φ, and of $(1-P_{\beta g})$ for the indices of valence states not belonging to the configuration Φ.

The total wave function Ψ can now be written as a sum over all possible configurations Φ of localized valence states in the crystal. The terms in this sum are not mutually orthogonal, but for the discussion of that part of the Hamiltonian which refers to the valence states of one particular atom only, the lack of orthogonality may be temporarily forgotten.

The largest contributions to these intra-atomic energy terms, besides the kinetic energy and the potential energy in the field of the nucleus and the core electrons, come from the mutual Coulomb repulsion of the valence states (since the core states are always occupied, their repulsion provides merely a constant shift). Indeed all the Coulomb repulsion integrals between valence states contain one term which is the largest and identical in all of them.[6] It is the Coulomb repulsion between two spherically symmetric charge distributions, each with total charge equal to one electronic charge, and a radial dependence equal to that of the valence states. If ν of the valence states at a particular atom are occupied, this term in the Coulomb repulsion becomes simply $C\nu(\nu-1)/2$.

For some arbitrary wave function Ψ, the number ν of occupied valence states at some particular atom is, in general, not fixed. Ψ may contain configurations Φ

[6] John C. Slater, *Quantum Theory of Atomic Structure* (McGraw-Hill Book Company, Inc., New York, 1960), Vol. I, pp. 311, 490.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which have different numbers of occupied valence states at the particular atom under examination. To the approximation where the configurations Φ are considered mutually orthogonal, a probability p_ν can be defined for Ψ, to have exactly ν occupied valence states at the particular atom. p_ν is simply the sum of the squares of all those configurations Φ which occur in the expansion of Ψ and have exactly ν occupied valence states at the particular atom.

Since the Coulomb repulsion $C\nu(\nu-1)/2$ is the largest among the many two-particle terms, it is important for the wave function Ψ to minimize the quantity

$$\tfrac{1}{2}\sum \nu(\nu-1)p_\nu=q, \tag{21}$$

with the summation going over the possible occupation numbers of the valence states, e.g., ν goes from 0 to 10 for the entire set of $3d$ states. There are two obvious conditions to be satisfied for the set of probabilities p_ν. The normalization condition is

$$\sum p_\nu=1, \tag{22}$$

and the total number of valence electrons is given by, say, $\bar{n}$ per atom on the average, so that

$$\sum \nu p_\nu=\bar{n}. \tag{23}$$

The solution to this simple minimization problem is found most easily in the following way. $\bar{n}$ is assumed to lie between two integers μ and $\mu+1$, i.e.,

$$\mu<\bar{n}<\mu+1. \tag{24}$$

q can then be expressed with the help of (22) and (23) as

$$q=(\mu/2)(2\bar{n}-\mu-1)+\tfrac{1}{2}\sum(\mu-\nu)(\mu+1-\nu)p_\nu. \tag{25}$$

The coefficients of p_μ and $p_{\mu+1}$ vanish, whereas the coefficients of all the other p's are positive. Therefore, q reaches its minimum for the values

$$p_\mu=\mu+1-\bar{n},\quad p_{\mu+1}=\bar{n}-\mu,$$
$$p_\nu=0 \quad \text{for} \quad \nu\neq\mu \quad \text{and} \quad \nu\neq\mu+1. \tag{26}$$

The values of p_μ and $p_{\mu+1}$ are obtained from (22) and (23) after all the other p_ν's have been set equal to zero.

The statement expressed in (26) has some striking consequences. For example, in the case of iron we have $\mu=7$, and it follows that any reasonable wave function Ψ for the whole iron crystal should contain only configurations which have either 7 or 8 occupied valence states on each atom. Similarly, in Ni we have $\mu=9$ and there ought not to be any configuration with less than 9 valence ($3d$) electrons at any lattice site.

The requirement (26) is generally violated by the itinerant electron model. Let us consider for example the Ni-Cu alloy with $100p$ atomic percent Cu substituted for Ni. We shall assume that $0.6-p$ holes per atom have to be distributed equally over two degenerate valence bands in the ferromagnetic ground state, and over four degenerate valence bands in the nonferromagnetic ground state. In terms of the numbers ν of holes we then have

$$p_\nu=\binom{2}{\nu}\left(1-\frac{0.6-p}{2}\right)^\nu\left(1+\frac{0.6-p}{2}\right)^{2-\nu} \text{ ferromagnetic,}$$

$$p_\nu=\binom{4}{\nu}\left(1-\frac{0.6-p}{4}\right)^\nu\left(1+\frac{0.6-p}{4}\right)^{4-\nu} \text{ nonferromagnetic.}$$

The second term in (25) becomes equal to $(0.6-p)^2/4$ for the ferromagnetic and $3(0.6-p)^2/8$ for the nonferromagnetic case.

The Coulomb repulsion C can be estimated as follows. The outermost $3d$ electrons see at least 1.6 unscreened nuclear charges. But since there are at least two of these outermost $3d$ orbits, the average unscreened nuclear charge in these orbits must be at least 2 so that such an outermost $3d$ orbit can be thought of as a hydrogenic $3d$ orbit in the field of an He nucleus. If we think of one of the two mutually repelling electronic charge distributions as completely concentrated at the nucleus (this tends to overestimate the value of C, but the hydrogenic $3d$ orbit has been chosen rather conservatively), C becomes equal to 4/9 Ry. The Coulomb repulsion energy is, therefore, $(0.6-p)/9$ Ry per hole for the ferromagnetic and $(0.6-p)/6$ Ry per hole for the nonferromagnetic case.

The surplus of Coulomb-repulsion energy for pure Ni, i.e., $p=0$, corresponds, according to our rough estimate, to about 10 000°K in the ferromagnetic and to about 15 000°K in the nonferromagnetic itinerant model, whereas the Curie temperature in Ni is found to be 650°K. It seems futile to discuss the itinerant model further without a serious attempt to find a modified wave function which is not beset by the large surplus of Coulomb repulsion in either the ferromagnetic or the nonferromagnetic state.

Since there are many configurations Φ with not more than $\mu+1$ and not less than μ occupied valence states on any lattice site, a good wave function Ψ for the whole crystal is expected to be a linear combination of all such configurations. Indeed any configuration Φ by itself does not give a particularly low expectation value for the kinetic energy and crystal potential terms in the Hamiltonian which correspond to a single electron moving in an effective potential of the lattice symmetry. Exactly these last terms were minimized by the itinerant model. They constitute the first order if the terms in the Hamiltonian are arranged in ascending powers of the overlap R between valence orbits on neighboring sites. The Coulomb repulsion C belongs to the zero order in such an expansion.

The basic idea is then to enforce the result (26) upon the itinerant model wave function by eliminating those configurations which violate the requirement (26) at some particular lattice site. Such a correlation corresponds exactly to the correlation in the Heitler-London wave function for the hydrogen molecule. But since we have now to do with $3d$ orbits the requirement (26) still

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

leaves many configurations at each lattice site, so that some additional correlation beyond the Heitler-London type (26) seems indicated. In particular Hund's rule should now be enforced at each lattice site.

4. A CORRELATED WAVE FUNCTION FOR THE VALENCE ELECTRONS

The ground state Ψ_0 in the itinerant electron model is obtained from the vacuum state Φ_0 by creating N electrons in Bloch states $\psi_{\kappa k}(x)$. Therefore, we have

$$\Psi_0 = \prod_{(\kappa k)} c^\dagger_{\kappa k} \Phi_0 , \tag{27}$$

where the set (κk) of occupied Bloch states is chosen in some manner, e.g., such as to minimize the expectation value of the total energy. The particular choice of the Bloch functions $\psi_{\kappa k}$ and the set (κk) is not under discussion right now, and we shall assume that both the functions $\psi_{\kappa k}(x)$ and the set (κk) have been determined once for all.

Ψ_0 is now expanded into configurations Φ of localized valence states by the method described at the beginning of the last section. The new (correlated) wave function Ψ is then obtained by transforming each configuration $\Phi(\alpha_1 f_1, \cdots, \alpha_M f_M)$ in the expansion of Ψ_0 into a linear combination of related configurations $\Phi(\gamma_1 h_1, \cdots, \gamma_M f_M)$. Most generally this transformation is described by an operator whose matrix elements give the coefficient of the configuration $\Phi(\gamma_1 h_1, \cdots, \gamma_M h_M)$ in the transformation of the configuration $\Phi(\alpha_1 f_1, \cdots, \alpha_M f_M)$.

A more specific assumption has to be made. As the most stringent assumption we will assume that the above operator O does not transfer electrons from one lattice site to another. Then it is helpful to simplify matters by assuming O to effect transfers of electrons within one atom independently of the transfers of electrons in neighboring atoms. We have, therefore, a set of operators $O^{(0)}$, $O^{(1)}$, $O^{(2)}$, $\cdots$ applicable to lattice sites which are occupied by 0, 1, 2, $\cdots$ electrons in the valence shell. The determination of these operators O is essentially equivalent to bringing about the necessary correlation of the valence states. In terms of the creation and annihilation operators we can write

$$O^{(\nu)} = \sum_{(\gamma_1 \cdots \gamma_\nu)} \sum_{(\alpha_1 \cdots \alpha_\nu)} \langle \gamma_1 \cdots \gamma_\nu | O^{(\nu)} | \alpha_1 \cdots \alpha_\nu \rangle$$
$$\times a^\dagger_{\gamma_\nu} \cdots a^\dagger_{\gamma_1} b_{\alpha_1} \cdots b_{\alpha_\nu} (1 - P_{\alpha_{\nu+1}}) \cdots (1 - P_{\alpha_{10}}) , \tag{28}$$

to be summed over all different groups of ν indices $(\alpha_1 \cdots \alpha_\nu)$ and $(\gamma_1 \cdots \gamma_\nu)$. In the case of $3d$ bands there is a maximum of 10 different indices.

For example the requirement (26) is enforced by setting all $O^{(\nu)} = 0$, except $O^{(\mu)} = O^{(\mu+1)} = 1$. This constitutes the generalization of the Heitler-London model for the hydrogen molecule. A somewhat weaker correlation is brought about by allowing the $O^{(\nu)}$ to deviate from the above value, but keeping each $O^{(\nu)}$ a multiple of the unit matrix.

It is now evident how Hund's rule can be favored. The configuration at each lattice site can be expanded into the states corresponding to the various spectral levels. The operators O can now be chosen so as to eliminate certain unfavorable spectral levels, or at least to reduce their contributions, whereas the levels which are lowest according to Hund's rule can be left unreduced. Therefore, the operators O may have off-diagonal elements. The values of the matrix elements of O can be determined variationally.

5. ALTERNATIVE FORMULATION FOR THE CORRELATED WAVE FUNCTION

If one is interested in a metal with only few electrons in the valence band, it is natural to expand all the results in powers of the density of valence electrons in the lattice. The operator O of the previous section is then best replaced by a slightly different expression which will be defined in this section.

The operator O will be expressed in terms of matrix elements $\vartheta_0 = 1, \vartheta_1(\gamma,\alpha), \vartheta_2(\gamma_1\gamma_2,\alpha_1\alpha_2)$, etc. These matrix elements are numbers out of which $\langle \gamma_1 \cdots \gamma_\nu | O^{(\nu)} | \alpha_1 \cdots \alpha_\nu \rangle$ can be computed by taking all kinds of combination of $\vartheta_0, \vartheta_1, \vartheta_2, \cdots$ and summing them.

First there is a summation over ν from 0 to 10 and for each ν a summation over all different ν-tuples $(\alpha_1 \cdots \alpha_\nu)$. Second, for each ν there is a summation over all partitions $P^{(\nu)}$ of the integer ν. Third, for each partition there is a summation over all possible groupings of the ν-tuple $(\alpha_1 \cdots \alpha_\nu)$ according to P, and over all permutations of $(\gamma_1 \cdots \gamma_\nu)$. A particular grouping of $(\alpha_1 \cdots \alpha_\nu)$ is written as $(\alpha^1 \cdots \alpha^\nu)$ where $\alpha^1 \cdots \alpha^\mu$ are associated with the 1's of the partition P, $(\alpha^{\mu_1+1}\alpha^{\mu_1+2})$, $(\alpha^{\mu_1+3}\alpha^{\mu_1+4})$, $\cdots$ with the 2's, etc. The permutation of $(\gamma_1 \cdots \gamma_\nu)$ is written $(\gamma^1 \cdots \gamma^\nu)$. With these preliminaries one can now write

$$O = -\frac{1}{\omega} \sum_{\nu=0}^{10} \sum_{(\alpha_1 \cdots \alpha_\nu)} \sum_{(\gamma_1 \cdots \gamma_\nu)} \sum_{P^{(\nu)}} \sum_{\substack{\text{groupings} \\ \text{of } (\alpha_a \cdots \alpha_\nu)}} \sum_{\substack{\text{permutations} \\ \text{of } (\gamma_1 \cdots \gamma_2)}}$$
$$\times \vartheta_1(\gamma^1\alpha^1)\vartheta_1(\gamma^2\alpha^2) \cdots \vartheta_2(\gamma^{\mu_1+1}\gamma^{\mu_2+1}\alpha^{\mu_1+1}\alpha^{\mu_1+2}) \cdots$$
$$\times a^\dagger_{\gamma_\nu} \cdots a^\dagger_{\gamma_1} b_{\alpha_1} \cdots b_{\alpha_\nu} . \tag{29}$$

The term $\nu = 0$ is put equal to 1.

The factor $1/\omega$ expresses the fact that the same wave function Ψ results, if the left-hand and the right-hand sides of (29) are just multiples of each other with the same factor for all operators.

Since the operators $O^{(\nu)}$ have an intuitive interpretation according to the preceding section, the problem arises of finding the matrix elements ϑ from known operators O by solving Eq. (29). The special case of the O's being chosen according to the requirement (26) is treated in the Appendix A. We shall assume that $\omega = 1$ from now on. This implies that $O^{(0)} = 1$, or in other words, that a configuration will not be discriminated against for having no electrons at all at some particular lattice site. Such an assumption is only reasonable if

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

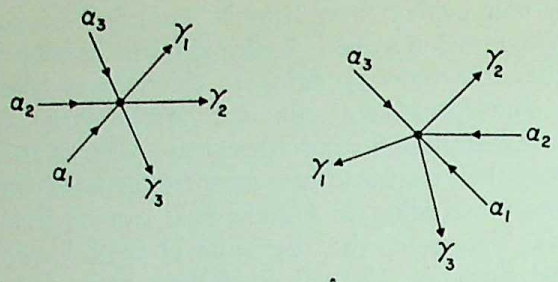


FIG. 1. Two equivalent vertices corresponding to ϑ_3.

there are fewer valence electrons on the average in the lattice than lattice sites.

The theory has been phrased in terms of electrons, but in the case of almost-filled bands it is more appropriate to consider the holes in the valence bands. This is done simply by switching the positions of $a^\dagger$ with respect to b in (28) and (29). The projection operator $1-P=1-a^\dagger b=ba^\dagger$ appearing in (28) then goes into $a^\dagger b=P$, which tests the absence of a hole exactly as $1-P$ tested the absence of an electron. Again, an equivalence between (28) and (29) can be established only in the cases where there is no physical reason to discriminate against a configuration for having no holes at all at some particular lattice site. Since the present investigation is directed mostly toward the understanding of the Ni-Cu alloys, all the following calculations will be done for the case of an almost-full valence band.

A picture can be associated with any particular term in (29), if each factor ϑ_ν is represented by a heavy dot, called a vertex, with ν directed lines across. The endpoints of the lines are labeled on the incoming end by α, and on the outgoing end by γ. A combination of such representations of ϑ's is called a picture. Two pictures are called equivalent if they can be brought to congruence including the labeling of lines, by shifting their parts. Figure 1 shows two equivalent vertices.

Conversely, given a picture, i.e., a combination of vertices with labeled ingoing and outgoing lines, one associates a product of ϑ's and operators $a^\dagger$, b with it. To each vertex corresponds a factor $\nu!\vartheta_\nu$ with the arguments occurring in the picture. Formula (29) can then be expressed as a sum over all pictures to be made with two sets of labels $(\alpha_1\cdots\alpha_{10})$ and $(\gamma_1\cdots\gamma_{10})$.

6. CALCULATION OF EXPECTATION VALUES

The correlated wave function Ψ is now given by

$$\Psi=\prod_g O_g\Psi_0, \tag{30}$$

where O_g is the operator (29) belonging to the lattice site g. Since Ψ_0 represents almost filled valence bands, $a^\dagger$ is replaced formally by b everywhere, and vice versa, everything else in (29) remaining at its place. The first thing to be done is the computation of the absolute value $(\Psi|\Psi)$ of the correlated wave function Ψ.

Formula (29) is inserted in (30) by attaching an additional index g to the ϑ_μ, b_γ, and $a^\dagger_\alpha$. The product over the lattice sites is worked out for both Ψ and Ψ^*, the scalar product is taken term by term, and the results are added up. Each term in this last sum can be represented by a picture in the following way:

To each $\vartheta_{\mu g}$ or $\vartheta^*_{\mu g}$ associate a heavy dot called a vertex, to be labeled by the lattice site g. Each operator $a^\dagger_{\alpha g}$ associated with $\vartheta_{\mu g}$ is represented by a solid line with an arrow going into the vertex of $\vartheta_{\mu g}$, and each operator $a_{\alpha g}$ associated with $\vartheta^*_{\mu g}$ is represented by a solid line with an arrow coming out of the vertex of $\vartheta^*_{\mu g}$. Each operator $b_{\gamma g}$ associated with $\vartheta_{\mu g}$ is represented by a dotted line leaving the vertex of $\vartheta_{\mu g}$, and the operator $b^\dagger_{\gamma g}$ associated with $\vartheta^*_{\mu g}$ is represented by a dotted line going into the vertex of $\vartheta^*_{\mu g}$. The lines are labeled exactly as the operators they represent. Figure 2 is an example of such a picture representing a term in $(\Psi|\Psi)$. Two pictures like Fig. 2 are called identical if they can be brought to congruence including labels and directions by merely shifting the parts of the picture without ever detaching a line from its vertex. If one permutes the arguments $\gamma_1\cdots\gamma_\mu$ in $\vartheta_{\mu g}$ the same picture arises according to this definition.

Conversely, given a picture like Fig. 2, it is immediately possible to write down the corresponding term in the evaluation of $(\Psi|\Psi)$. For each vertex with μ incoming solid lines one writes an operator

$$\mu!\vartheta_{\mu g}(\gamma_1\cdots\gamma_\mu\alpha_1\cdots\alpha_\mu)b_{\gamma_1 g}\cdots b_{\gamma_{10}g}a^\dagger_{\alpha_1 g}\cdots a^\dagger_{\alpha_\mu g}$$

with the corresponding labels, and analogously for each vertex with μ outgoing solid lines. All operators belonging to a ϑ^* are written to the left of all operators belonging to a ϑ. The order of operators belonging to different lattice sites is immaterial. Operators belonging to the same lattice site have to be pulled together in the order suggested by (29). A particular label βg on $a^\dagger$ can then not occur more than once, and similarly any particular label on b, $b^\dagger$, or a cannot occur more than once. The freedom left in ordering the operators belonging to different lattice sites cannot change the expectation value of the whole operator product because of the simple anticommutators (12) for $a^\dagger$ and b.

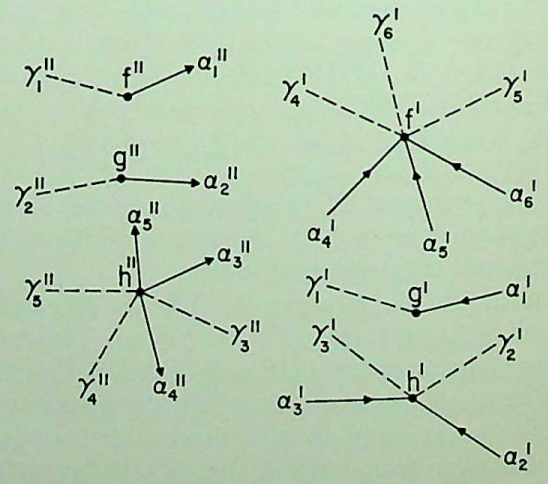


FIG. 2. Example of a picture for a term in (30).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The absolute value $(\Psi|\Psi)$ becomes now a sum over all different pictures like Fig. 2, where one associates an operator with each picture according to the rules of the previous paragraph, and then computes the expectation values of that operator with the wave function Ψ_0.

The computation of an expectation value like $(\Psi|b^\dagger_{\alpha f}b_{\gamma h}|\Psi)$ which arises from the one particle operator (16) is done in almost the same way. $b^\dagger_{\alpha f}$ is represented by a small cross labeled αf with an ingoing dotted line, and $b_{\gamma h}$ is represented by a small cross labeled γh with an outgoing dotted line, in addition to all the features in the pictures for $(\Psi|\Psi)$. Everything stays the same as with $(\Psi|\Psi)$ except the operators associated with ϑ^* have to be written to the left of $b^\dagger_{\alpha f}b_{\gamma h}$, and the operators associated with ϑ stand to the right. The small crosses associated with $b^\dagger_{\alpha f}$ and $b_{\gamma h}$ are called external points to distinguish them from the vertices (heavy dots).

7. EVALUATION IN TERMS OF DIAGRAMS

The expectation value of the operator associated with a picture like Fig. 2 can be evaluated for the uncorrelated wave function Ψ_0 with the help of Wick's theorem. According to this theorem one has to run through all complete systems of contractions, that is through all ways of grouping the annihilation and creation operators into pairs. The contribution of such a complete system of contractions is given by the product of the expectation values of each pair in the contraction for the wave function Ψ_0. The order of the operators in each pair has to be the same as in the original operator, and there is a sign to be determined from the number of permutations necessary in order to bring the partners of each pair together.

In the case of the uncorrelated wave function Ψ_0 only the pairs consisting of one annihilation and one creation operator give a nonvanishing result. There are four types of such pairs and their expectation values are easily computed with the help of the definitions (10) and (11), as well as the relations (5). They are

$$\begin{aligned}
&(\Psi_0|a_{\alpha f}a^\dagger_{\gamma h}|\Psi_0)=W(\alpha f,\gamma h),\\
&(\Psi_0|b_{\alpha f}a^\dagger_{\gamma h}|\Psi_0)=\sum_{\beta g}S(\alpha f,\beta g)W(\beta g,\gamma h),\\
&(\Psi_0|a_{\alpha f}b^\dagger_{\gamma h}|\Psi_0)=\sum_{\beta g}W(\alpha f,\beta g)S(\beta g,\gamma h),\qquad (31)\\
&(\Psi_0|b^\dagger_{\alpha f}b_{\gamma h}|\Psi_0)=S(\gamma h,\alpha f)\\
&\quad-\sum_{\beta' g'\beta'' g''}S(\gamma h,\beta'' g'')W(\beta'' g'',\beta' g')S(\beta' g',\alpha f),
\end{aligned}$$

where W and S are defined by

$$\begin{aligned}
W(\alpha f,\gamma h)&=\frac{1}{L}\sum_{(\kappa k)}U^*_{\alpha f,\kappa k}U_{\gamma h,\kappa k},\\
S(\alpha f,\gamma h)&=\frac{1}{L}\sum_{\kappa k}V_{\kappa k,\alpha f}V_{\kappa k,\gamma h}.\qquad (32)
\end{aligned}$$

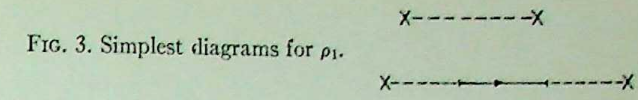


FIG. 3. Simplest diagrams for ρ_1.

The summation over (κk) means that only those indices κk are included in the summation which correspond to empty Bloch states of Ψ_0. It is easily checked that the matrix $S(\alpha f,\gamma h)$ is just the inverse of the overlap matrix $R(\alpha f,\gamma h)$ defined in (7).

It is immediately evident from writing down the operator corresponding to a picture like Fig. 2 that a and $a^\dagger$, as well as $b^\dagger$ and b, have always to be paired in the sequence in which they appear in the first and last formula of (31). However, the pairs made up of b and $a^\dagger$, or of a and $b^\dagger$, may occur in either sequence, except if the two partners of such a pair belong to the same lattice site. In that case b always precedes $a^\dagger$, and a precedes $b^\dagger$, exactly as in the second and third formula of (31). If b and $a^\dagger$, or a and $b^\dagger$ do not belong to the same lattice site, the commutation relations (12) tell us that only the sign of the expectation value changes when the order of the partners in the pair is changed. But this sign change is accounted for when the over-all sign to be attached to the product of the quantities (31) is determined.

A complete system of contractions can be represented with the help of pictures like Fig. 2 in the following way. A pair of the type $aa^\dagger$ is represented by simply connecting the two solid lines representing a and $a^\dagger$, the arrow points from a to $a^\dagger$. Similarly a pair of type $ba^\dagger$ is represented by tying the dotted line of b to the solid line of $a^\dagger$, inserting a short cross line such as in Fig. 3 at the point where the dotted and the solid lines meet. This short cross line marks the end of the dotted and the beginning of the solid line, and corresponds to the summation over βg which appears in the formula (31) for $(\Psi_0|b_{\alpha f}a^\dagger_{\gamma h}|\Psi_0)$. A similar treatment is given to a pair $ab^\dagger$. Finally, since a pair $b^\dagger b$ gives rise to two terms as shown in (31), it will be represented in two ways as shown in Fig. 3, first by simply connecting the two dotted lines of $b^\dagger$ and b, and second by joining the dotted line of $b^\dagger$ with the dotted line of b through the intermediary of a directed solid line. The arrows go from b to $b^\dagger$ which is reflected in the inverted arguments of the last formula in (31) as compared to the preceding formulas in (31). The minus sign appearing in the last part of $(\Psi_0|b^\dagger_{\alpha f}b_{\gamma h}|\Psi_0)$ will be absorbed in the general rules determining the overall sign. This pictorial representation of a particular complete system of contractions by connecting the appropriate lines in a picture like Fig. 2 is called the diagram associated with that system of contractions. Two diagrams are identical if they can be brought into complete congruence (including labels and arrows) by shifting the lines in the drawing without ever breaking any of them.

In the present formulation the particle lines (solid or dotted) run from one vertex to another, but they do not

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

run through a vertex. None of the solid lines at some vertex is in a particular relationship with any given dotted line belonging to the same vertex, and vice versa. However, if one starts from a given diagram in order to write its contribution to $(\Psi|\Psi)$, the result depends on the order in which the arguments α and γ are listed in the ϑ corresponding to some particular vertex. In establishing the over-all sign to be attached to the contribution, it is convenient to adopt the following rule: The particle line ending in α_i at some vertex ϑ_μ continues on the other side with the particle line labeled γ_i, and the particle line ending in γ_j in ϑ^*_ν continues with α_j. In this manner all particle lines are joined two by two to form so-called loops. A closed loop closes itself in the diagram, an open loop ends at external points, which are the ones represented by small crosses. The over-all sign to be associated with a particular diagram is obtained in the following manner. The solid and dotted lines are organized into loops according to the rules of the previous paragraphs. A factor (-1) is now inserted for each segment of solid line in the diagram and for each closed loop.

In the original picture, e.g., in Fig. 2, the labels associated with outgoing solid lines were all different from one another; similarly, the labels of the incoming solid lines were all different; and similarly for the dotted lines, excepting the labels of dotted lines attached to external points. The expectation value associated with a picture like Fig. 2 is the sum over the contributions from all different diagrams to be drawn from the picture. The sum of these contributions is easily found to be an antisymmetric function of the labels associated with solid outgoing lines, of the labels associated with solid incoming lines, etc. As long as the over-all sign and the individual factors associated with a diagram are determined by the above rules, using in particular only the pair expectation values (31), it is formally possible to let labels of outgoing solid lines coincide, or to let labels of incoming solid lines coincide, etc. The sum over all diagrams associated with a picture containing identical labels on outgoing solid lines, or on incoming solid lines, etc., will give a vanishing result.

Therefore, it is now possible to let all the labels in a diagram run through all admissible values, independently of one another. The restriction to admissible values refers to the fact that only localized valence states are involved, but no conduction states. The overcounting resulting from this mutual independence makes it necessary to divide by

$$\mu_1!\mu_2!\cdots\nu_1!\nu_2!\cdots(1!)^{\nu_1+\mu_1}(2!)^{\nu_2+\mu_2}\cdots$$

where $\mu_1,\mu_2,\cdots$ is the number of $\vartheta_1,\vartheta_2,\cdots$, and $\nu_1,\nu_2,\cdots$ is the number of $\vartheta^*_1,\vartheta^*_2,\cdots$ in the diagram.

In summary, the expectation values of Ψ can be written as a sum over all different diagrams. A diagram consists of vertices, solid lines, dotted lines, and possibly external points. There are two kinds of vertices, with outgoing solid (incoming dotted) lines and with incoming solid (outgoing dotted) lines. There are two kinds of external points, with one incoming dotted line, and with one outgoing dotted line. Each line going out of one vertex (or external point) has to be tied to a line going into some vertex (or external point). A solid line is tied directly to another solid line; a solid line is tied to a dotted line, and vice versa, by inserting a short cross bar at the joining point; a dotted line can be tied either directly to another dotted line, or indirectly through the intermediary of a solid line segment, as shown in Fig. 3. Each vertex is labeled as a lattice site, and the lines going into and coming out of a vertex carry labels of band indices for localized valence states. Each external point carries the label of a lattice site with band index. The lattice site labels on the vertices and the external points are all different from one another. Similarly the band index labels on any particular vertex are different from one another. Two diagrams are different, if they cannot be brought to coincidence including the direction of lines and the labels on them by simply shifting the vertices on the paper without breaking any lines.

The contribution of a diagram is obtained as follows:

(1) To each vertex associate a factor ϑ_μ or ϑ^*_μ where μ is the number of incoming or outgoing solid lines;

(2) Attach a different lable βg to each short cross bar in the diagram;

(3) With each solid line from αf to γh associate a factor $W(\alpha f,\gamma h)$;

(4) With each dotted line from αf to γh associate a factor $S(\alpha f,\gamma h)$;

(5) Insert a factor (-1) for each solid line and each closed loop in the diagram;

(6) Insert a factor

$$1/\mu_1!\mu_2!\cdots\nu_1!\nu_2!\cdots(1!)^{\nu_1+\mu_2}(2!)^{\nu_2+\mu_2}\cdots,$$

where μ_1 is the number of ϑ_1,μ_2 the number of ϑ_2, etc., ν_1 the number of ϑ^*_1, etc. in the diagram;

(7) Sum independently over all labels excepting naturally the labels of external points, the labels associated with vertices run only over the band indices of localized valence states, whereas the labels associated with short cross bars run over all band indices (valence and conduction).

A special rule applies to the diagrams with more than one pair of external points, such as occur in the computation of Q_2 in (18) which involves the operators $b^\dagger_{\alpha_2 f_2}b^\dagger_{\alpha_1 f_1}b_{\gamma_1 h_1}b_{\gamma_2 h_2}$. If the open loops go from $\gamma_1 h_1$ to $\alpha_2 f_2$ (rather than $\alpha_1 f_1$) and from $\gamma_2 h_2$ to $\alpha_1 f_1$ (rather than $\alpha_2 f_2$) an extra factor (-1) has to be inserted.

In the following discussion of particular diagrams it is convenient to distinguish various types of labels which appear in the diagrams by the following rules: The double label consisting of a small Greek letter for the band index and a small Latin letter for the lattice site is replaced by one capital letter. An external point with outgoing dotted line is labeled Γ', and one with

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

incoming dotted line is labeled Γ''. An outgoing dotted line on a vertex is labeled H', and an incoming dotted line on a vertex is labeled H''. A short cross bar is labeled G' when it joins an incoming solid line to an outgoing dotted line, and it is labeled G'' when it joins an outgoing solid line to an incoming dotted line. Finally, an incoming solid line on a vertex is labeled F' and an outgoing solid line on a vertex is labeled F''. Clearly, the labels Γ' and Γ'' are fixed, the labels F', F'', H', and H'' run through all localized valence states in the lattice, whereas the labels G' and G'' run through all localized states (valence and conduction) in the lattice. Also the labels Γ' and Γ'' may stand for valence as well as for conduction states.

8. LINKED DIAGRAMS

Any part of a diagram which can be completely separated from the remaining diagram without breaking any line, is called an unlinked part. Once a diagram is divided up completely into unlinked parts, its total contribution can easily be obtained from calculating the separate contributions for each unlinked part, the total being equal to the product of its parts.

Unlinked parts without external points are called closed linked diagrams, if they cannot be broken up further into unlinked parts. The contributions of all closed linked diagrams appear simply as a factor

$$\exp[\sum_{\text{closed linked}} \Lambda] \tag{33}$$

to the diagrams with external lines. Λ stands for the number calculated, according to the rules of the previous section, for a particular closed linked diagram.

Since (33) is just the normalization integral (Ψ,Ψ), this factor divides out if ρ_1 and ρ_2 are computed with a normalized correlated wave function Ψ. Therefore, we can write with the help of (16) and (18)

$$\rho_1(x,y)=\sum_{\Gamma'\Gamma''}\varphi_{\Gamma'}(x)\varphi^*_{\Gamma''}(y) \times \sum_{\substack{\text{open linked with}\\ 2\text{ external points}}} \Lambda(\Gamma'\Gamma''), \tag{34}$$

$$\rho_2(\xi,x;\eta,y)=\tfrac{1}{2}\rho_1(\xi,\eta)\rho_1(x,y)-\rho_1(\xi,y)\rho_1(x,\eta) +\tfrac{1}{2}\sum_{(\Gamma'_1\Gamma'_2)(\Gamma''_1\Gamma''_2)}\begin{vmatrix}\varphi_{\Gamma'_1}(\xi) & \varphi_{\Gamma'_1}(x)\\ \varphi_{\Gamma'_2}(\xi) & \varphi_{\Gamma'_2}(x)\end{vmatrix} \times\begin{vmatrix}\varphi^*_{\Gamma''_1}(\eta) & \varphi^*_{\Gamma_2''}(y)\\ \varphi^*_{\Gamma''_1}(\eta) & \varphi^*_{\Gamma_2''}(y)\end{vmatrix} \times \sum_{\substack{\text{open linked with}\\ 4\text{ external points}}} \Lambda(\Gamma'_1\Gamma'_2,\Gamma''_1\Gamma''_2). \tag{35}$$

The interest of the last formula lies in the net separation of the "properly correlated" part of ρ_2 by the diagram analysis, i.e., the part of ρ_2 which cannot be simply understood as the effect of many-particle statistics and the exclusion principle.

According to the last two formulas it is only necessary to obtain the contributions of all open linked diagrams with two or four external points. These diagrams can be obtained in a systematic fashion from the closed linked diagrams, of which there are fewer, of course.

A closed linked diagram can be associated to each open linked diagram by simply joining the end points of each open loop, i.e., joining Γ' and Γ'', Γ'_1 and Γ''_1, Γ'_2 and Γ''_2. In doing this, a diagram may arise which is not allowed under previously established rules. Indeed, the external points are always reached by a dotted line, but these dotted lines may not be attached at their other ends to a point H' of H''. Instead, they may both be attached to a point G' or G''. The closing of such an open loop would then lead to a sequence solid-dotted-solid line where neither end of the dotted line is attached to a point H' or H'', contrary to our rules. However, in view of (5) and (32), such an illegal sequence may be contracted into one solid line, and in this way a legitimate closed linked diagram is obtained.

Conversely, it is evident that we obtain all open linked diagrams if we open in turn every solid or dotted line, or every couple of solid and/or dotted lines in all closed linked diagrams, provided the diagram does not decay into two unlinked parts. The rule for opening solid or dotted lines is simple. First write the expression for the closed linked diagram. Second, in order to obtain the contribution to ρ_1 make in turn the following replacements.

$$W(HF)\rightarrow\sum_{G'_0G''_0}W(HG'_0)S(G'_0\Gamma'')S(\Gamma'G''_0)W(G''_0F),$$
$$S(HF)\rightarrow -S(H\Gamma'')S(\Gamma'F), \tag{36}$$

for every factor W and S. Third, in order to obtain the contribution to the properly correlated part of ρ_2 make the replacements (36) in turn on every couple of factors with the pairs of external points (Γ''_1,Γ'_1) and (Γ''_2,Γ'_2). The over-all sign remains the same as in the closed linked diagram.

9. THE FIRST-ORDER DENSITY FUNCTION

The original expression for ρ_1 contains two terms which cannot be associated with a closed linked diagram. Their diagrams are given in Fig. 3, and their contributions to ρ_1 follow directly from the last formula (31) together with (3) and (11)

$$\sum_{[\kappa k]}\psi_{\kappa k}(x)\psi^*_{\kappa k}(y), \tag{37}$$

where $[\kappa k]$ is the set of Bloch states which are occupied by electrons. The complementary set, occupied by holes, is designated by (κk). This is just the ρ_1 from the uncorrelated wave function Ψ_0 which forms the base for the itinerant electron model.

It is of interest to notice that the total number of electrons in the lattice

$$N=\int dx\rho_1(x,x) \tag{38}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

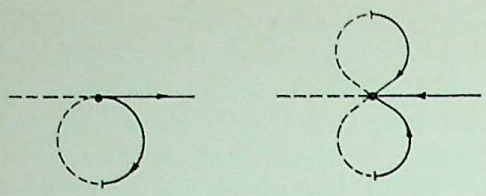


FIG. 4. Examples of bubbles.

is correctly given by (37). The correction terms which come from more complicated diagrams have to cancel one another in the integral (38). This fact is easily seen from (36) if we note that the integral (38) for any open linked diagram with two external points is equal to (± 1) times the contribution of the corresponding closed linked diagram; $+1$, if the open diagram resulted from opening a solid line; -1, if the open diagram resulted from opening a dotted line. A contribution to N from opening a solid line in a closed linked diagram is exactly canceled by the contribution to N from opening a dotted line in the same closed linked diagram.

This last property of the diagrams enables us to order the contributions to ρ_1 according to powers in the density m of holes in the valence band. m is defined as the ratio M/L where M is the average number of holes in localized valence states for the whole lattice. The order of some diagram is simply the power of m which occurs in the contribution of the associated closed linked diagram.

There are more terms of order one than (37). They arise from the closed linked diagrams which contain only ϑ_1, but no ϑ_2, ϑ_3, etc. Due to the simple structure of these diagrams they can all be evaluated. Their contribution together with (37) leads to substituting each occupied state $\psi_{\kappa k}(x)$ by a state $\bar{\psi}_{\kappa k}(x)$.

It is shown in Appendix B, how taking account of ϑ_1 throughout all possible diagrams is tantamount to redefining W in terms of $\bar{\psi}_{\kappa k}$ rather than $\psi_{\kappa k}$ and to a proper modification of the operators $\vartheta_2, \vartheta_3, \cdots$.

Another class of simple diagrams for ρ_1 is provided by the opening of the dotted and the solid lines in one of the two partial diagrams in Fig. 4. The vertices to which these "bubbles" are attached have other ingoing and outgoing lines, but the additions implied in all those other lines can be performed even before opening the bubbles. The result appears to the bubbles exactly as ϑ_1 would have appeared to them. Quite generally, any part of a diagram which can be reduced to ϑ_1 by taking loops off a vertex, appears to that diagram like a ϑ_1 insertion. Therefore, the effect of all such "dangling" parts in a diagram can be treated exactly as the effect of ϑ_1 contributions was treated in Appendix B.

The following result emerges: The treatment of ϑ_1 insertions is extended to include the "dangling" parts of a diagram. All that appears in the redefinition of Appendix B is the sum of ϑ_1 and all possible dangling parts. Since $\vartheta_1, \vartheta_2, \vartheta_3, \cdots$ are essentially determined by the assumptions concerning $O^{(1)}, O^{(2)}, O^{(3)} \cdots$ it seems that one is not free to choose ϑ_1 so as to cancel all the dangling parts. But the freedom to choose Ψ_0 still remains, and one might determine $\vartheta_1, \vartheta_2, \vartheta_3, \cdots$ variationally in order to determine $O^{(1)}, O^{(2)}, O^{(3)}, \cdots$ afterwards. In that case it is just as well to take $\vartheta_1 = 0$ and neglect all diagrams with dangling parts. We shall adopt this last viewpoint in the future development.

Up to the second order in the density of holes there is only one closed linked diagram out of which contributions to ρ_1 arise. This diagram is given in Fig. 5. It is of second order in ϑ_2. With assumed values for the constants determining the band structure, and for the overlap integrals, the band-structure term in the expectation value of the energy becomes, therefore, quadratic in ϑ_2. The Coulomb-repulsion term related to these second-order terms in ρ_1 is not only quartic in ϑ_2, but also in the density of holes, so that it can be neglected with respect to the band-structure term arising from ρ_1.

The explicit calculation of ρ_1 is done most easily and most usefully in terms of the Fourier transforms with respect to the lattice sites. Since ϑ_2 differs from zero only when all of its arguments are at the same lattice site, the Fourier transform does not affect ϑ_2 which simply becomes independent of the momenta exchanged between its in- and outgoing dotted and solid lines. Since the band indices are all carried along explicitly, the wave vectors k are always to be reduced to the first Brillouin zone. This implies conservation of wave vectors at each vertex up to a vector of the reciprocal lattice.

The terms in ρ_1 corresponding to the diagrams of Fig. 5 can easily be written down if the overlap $R(\alpha f, \gamma h)$ is simplified to $\delta_{\alpha\gamma}\delta_{fh}$. The localized orbits are then effectively Wannier states, and the two vertices of the diagrams in Fig. 5 have to be at the same lattice site. The particle propagator $W(F''F')$ has to be evaluated at the same lattice site $f' = f''$. It is shown in Appendix C that these assumptions have the following consequences. W differs from zero only if the orbits with band indices α' and α'' belong to the same irreducible representation of the crystal symmetry operations around the fixed lattice site $f' = f''$. For subsets $\{\alpha'\}$ and $\{\alpha''\}$ of the same irreducible representation, W is a multiple of the unit matrix.

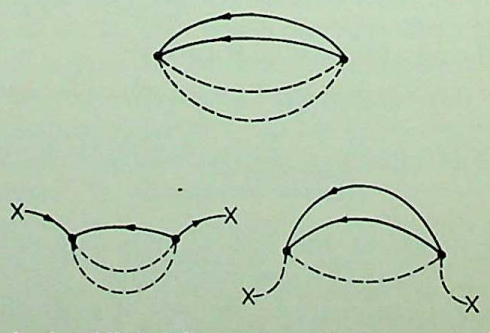


FIG. 5. Only closed linked diagram contributing to ρ_1 in the second order of density, and the two ρ_1 diagrams arising from it.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

As long as the spin variable is specifically excluded from transformation under the crystal symmetry operations, W may have nonvanishing elements between orbits of the same irreducible representation, but different spin directions. Such a situation is typical of an uncorrelated wave function Ψ_0 which describes an antiferromagnetic metal or one with some kind of spiral structure. Although such cases are not hard to handle, we shall simplify matters by writing down our formulas in their absence. It follows then that

$$W(\alpha'' f,\alpha' f)=m_\alpha \delta_{\alpha'\alpha''}, \tag{39}$$

where m_α is the probability of finding the orbit $\alpha=\alpha'=\alpha''$ at f occupied by a hole.

The Fourier transforms are written in terms of

$$u_{\beta\kappa}(k)=e^{ikg}U_{\beta g,\kappa k}. \tag{40}$$

According to (4) this quantity indicates the contribution of $\psi_{\kappa k}(x)$ to the localized states of index β. There is no simple relationship between κ and β. In the most general case the index κ simply numbers the consecutive Bloch functions belonging to the wave vector k. However, if the picture of a conduction band as somehow distinguishable from the valence bands is at all significant, the following assumptions about $u_{\beta\kappa}(k)$ can be made. For given β and κ the quantity $u_{\beta\kappa}(k)$ varies with the wave vector k fairly abruptly in such a manner that in most regions of the Brillouin zone $\psi_{\kappa k}(x)$ can be called either valence-like or conduction-like. Since $u_{\beta\kappa}(k)$ is assumed to be a continuous function of k, the transition regions have a nonvanishing but small volume in the reciprocal space.

The propagation function $W(\alpha f,\gamma h)$ of (32) is now written as

$$W(\alpha f,\gamma h)=\frac{1}{L}\sum_{(\kappa k)} u^*{}_{\gamma\kappa}(k)u_{\alpha\kappa}(k)e^{ik(f-h)}. \tag{41}$$

According to the above assumptions concerning $u_{\beta\kappa}(k)$ the propagation of a hole from αf to γh is done mainly by those terms κk in the sum which have the same character as α and γ. In particular, if α and γ have different character $W(\alpha f,\gamma h)$ is small, and if α and γ refer both to valence states $W(\alpha,f,\gamma h)$ does not exceed the density m of holes in valence states. It is this last statement which allows us to classify diagrams according to the number of solid lines in them.

The contribution of the diagrams in Fig. 5 to ρ_1

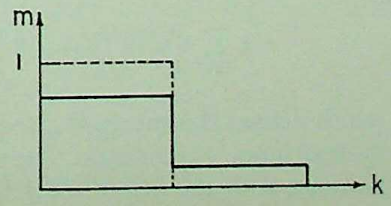


FIG. 6. Schematic distribution of holes in reciprocal space with (drawn out) and without (dotted) correlation.

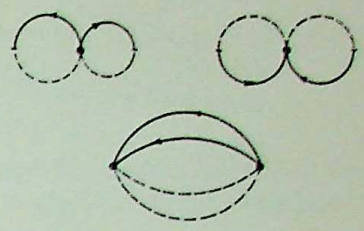


FIG. 7. Diagrams contributing to the properly correlated part of ρ_2 in the second order of density.

becomes

$$\frac{4}{L}\sum_{k(\kappa)(\lambda)}\sum_{\alpha\beta(\beta'\beta'')\gamma}\psi_{\kappa k}(x)u_{\alpha\kappa}(k)$$
$$\times\vartheta_2\begin{pmatrix}\alpha & \beta\\ \beta' & \beta''\end{pmatrix}m_\beta\vartheta_2^*\begin{pmatrix}\gamma & \beta\\ \beta' & \beta''\end{pmatrix}u_{\gamma\lambda}^*(k)\psi_{\lambda k}^*(y)$$
$$-\frac{4}{L}\sum_{k\kappa\lambda}\sum_{\alpha\beta(\beta'\beta'')\gamma}\psi_{\kappa k}(x)u_{\alpha\kappa}(k)$$
$$\times\vartheta_2^*\begin{pmatrix}\beta' & \beta''\\ \alpha & \beta\end{pmatrix}m_{\beta'}m_{\beta''}\vartheta_2\begin{pmatrix}\beta' & \beta''\\ \gamma & \beta\end{pmatrix}u_{\gamma\lambda}^*(k)\psi_{\lambda k}^*(y). \tag{42}$$

The summations (κ) and (λ) are restricted to the labels which are occupied by holes in Ψ_0, and the summation over $(\beta'\beta'')$ goes over all pairs of valence state indices. Since all indices α, β, β', β'', γ refer to valence states, only those indices κ and λ contribute to (42) which correspond to valence like states. The interpretation of (42) is therefore meant to apply only to valence states.

The first term in (42) describes a decrease in the occupation probability for those states which were originally fully occupied by valence holes. The second term spreads the holes which were freed by the first term over the whole available reciprocal space of valence character. The situation is schematically shown in Fig. 6. The kinetic energy and crystal-potential terms in the total energy can be obtained immediately.

10. THE SECOND-ORDER DENSITY FUNCTION

The closed linked diagrams which give contributions to the properly correlated part of ρ_2 have to have at least two solid and two dotted lines. On the other hand, as long as we are interested only in terms up to the second order in the density of holes, the only diagrams to be considered have not more than two solid lines. This shows that the only contributions of interest come from the diagrams in Fig. 7. Again these diagrams are at most of the second order in ϑ_2, so that the determination of ϑ_2 by minimizing the expectation value of the total energy for the correlated wave function Ψ is straightforward up to the second order in the density of holes.

The case of small overlap between valence states is of particular interest. The largest terms in the Hamiltonian are the Coulomb integrals with all orbits belonging to the same lattice site. The contribution of the diagrams above can then be evaluated quite easily,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

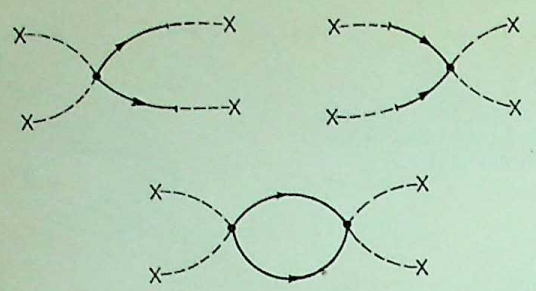


FIG. 8. The diagrams which contribute to the formula (43).

because in the lowest order of density and overlap only the diagrams in Fig. 8 with all external points at the same lattice site as that of ϑ_2 make a contribution.

The computation is straightforward for the same assumptions as were used in the derivation of (42). Together with the contribution of the valence hole term in (37) to the first term in (35) one finds for the terms in ρ_2 which are quadratic in the density

$$\bar{\rho}_2=\tfrac{1}{2}\sum_{(\alpha_1\alpha_2)(\gamma_1\gamma_2)}\begin{vmatrix}\varphi_{\alpha_1}(x) & \varphi_{\alpha_1}(\xi)\\ \varphi_{\alpha_2}(x) & \varphi_{\alpha_2}(\xi)\end{vmatrix}\begin{vmatrix}\varphi^*_{\gamma_1}(y) & \varphi^*_{\gamma_1}(\eta)\\ \varphi^*_{\gamma_2}(y) & \varphi^*_{\gamma_2}(\eta)\end{vmatrix}$$

$$\times\sum_{(\beta_1\beta_2)}\left[\delta\begin{pmatrix}\alpha_1 & \alpha_2\\ \beta_1 & \beta_2\end{pmatrix}+2\vartheta_2\begin{pmatrix}\alpha_1 & \alpha_2\\ \beta_1 & \beta_2\end{pmatrix}\right]m_{\beta_1}m_{\beta_2}$$

$$\times\left[\delta\begin{pmatrix}\gamma_1 & \gamma_2\\ \beta_1 & \beta_2\end{pmatrix}+2\vartheta^*_2\begin{pmatrix}\gamma_1 & \gamma_2\\ \beta_1 & \beta_2\end{pmatrix}\right] \quad (43)$$

at each lattice. The summations go over all different couples of valence states $(\alpha_1\alpha_2)$, $(\beta_1\beta_2)$, and $(\gamma_1\gamma_2)$. The function $\delta=1$ if the couples $(\alpha_1\alpha_2)$ and $(\beta_1\beta_2)$ are equal, otherwise $\delta=0$.

The function (41) is positive definite for $x=y$ and $\xi=\eta$. For given values of the densities m_β the ϑ_2 can be chosen such as to make $\bar{\rho}_2$ vanish. Such a requirement is reasonable if the one-center Coulomb and exchange integrals are much larger than the nearest neighbor overlaps divided by the atomic distance. This is just the assumption of narrow bands, and it leads immediately to a formula for ϑ_2, namely

$$2\vartheta_2\begin{pmatrix}\alpha_1 & \alpha_2\\ \gamma_1 & \gamma_2\end{pmatrix}=-\delta\begin{pmatrix}\alpha_1 & \alpha_2\\ \gamma_1 & \gamma_2\end{pmatrix}. \quad (44)$$

With this value for ϑ_2 one can compute the corrections to ρ_1 up to the second order in the densities m_β according to the preceding section. The redistribution of electrons in reciprocal space because of the strong correlation (44) is different according to the distribution of electrons in Ψ_0. Therefore, (44) already leads to a difference between ferromagnetic and nonferromagnetic arrangement of electrons in Ψ.

The advantage which correlation (44) gives to the ferromagnetic state, arises from having to overcome less accidental crowding of holes on the same site in Ψ_0. Correlation (44) forces fewer holes to be scattered outside of the Fermi surface if the holes all have the same spin direction. Correlation (44), however, prevents the electrons from taking advantage of the energy gain they experience when two of them crowd into the same lattice site with their spins parallel.

Quite generally some energy gain can be expected if the ϑ_2 as given in (44) is reduced by some small quantity ϵ,

$$2\vartheta_2\begin{pmatrix}\alpha_1 & \alpha_2\\ \gamma_1 & \gamma_2\end{pmatrix}=-\delta\begin{pmatrix}\alpha_1 & \alpha_2\\ \gamma_1 & \gamma_2\end{pmatrix}+\epsilon\begin{pmatrix}\alpha_1 & \alpha_2\\ \gamma_1 & \gamma_2\end{pmatrix}, \quad (45)$$

since the correlation energy connected with (43) increases with ϵ^2, whereas the single-particle energy due to the redistribution of electrons in reciprocal space decreases with ϵ. The coefficient of the former increase is of the order of the one-center Coulomb integral C, whereas the coefficient in the latter decrease is of the order of the valence bandwidth $\bar{E}$. Stabilization occurs at $\epsilon=\bar{E}/C$; the energy gain per lattice site is $\bar{E}^2/C$ times the density of holes squared.

The constant C to be inserted again depends on the spin arrangement in Ψ. The expression (44) has to be integrated over $x=y$ and $\xi=\eta$ with the kernel $e^2/|x-\xi|$. The resulting integral over the products of the 2 by 2 determinants is just equal to the Coulomb integral minus the exchange integral in the important case $(\alpha_1\alpha_2)=(\gamma_1\gamma_2)$. The exchange integral differs from zero only when α_1 and α_2 have the same spin, so that C is smaller in the ferromagnetic case, i.e., the energy gain from relaxing the correlation (42) is greater.

Among the off-diagonal elements $(\alpha_1\alpha_2)\neq(\gamma_1\gamma_2)$ some lead to exchange integrals (spin-flip transition). In the nonferromagnetic case, these off-diagonal elements may be expected to enforce Hund's rule locally to some extent.

11. TWOFOLD DEGENERATE VALENCE BAND

As the simplest possible application of the formulas (42) and (43), the case of a twofold degenerate band will be treated. This corresponds to the conditions in the Ni-Cu alloys, where the d states are in a cubic environment and there are some holes in the Bloch states made up of the x^2-y^2 and $2z^2-x^2-y^2$ states.

With the single particle Hamiltonian H_0 for the kinetic energy and the crystal potential one has $H_0\psi_{\kappa k}(x)=E_{\kappa k}\psi_{\kappa k}(x)$. In computing the expectation value of H_0 for the correlated wave function Ψ a quantity similar to the left-hand side of (39) arises. By the same group theoretical arguments as in Appendix C, it can be shown for valence states α and γ that

$$\frac{1}{L}\sum_{(k\kappa)}u^*_{\gamma\kappa}(k)E_{\kappa k}u_{\alpha\kappa}(k)=m_\alpha E_\alpha\delta_{\alpha\gamma}, \quad (46)$$

which defines the energy E_α, the average energy of the valence holes in the state α. The expectation value of H_0 becomes with the help of (42) for the correlated wave

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

function, omitting the contribution of the conduction-like states as being unaffected by the correlation,

$$\sum_{\alpha}(\bar{E}_\alpha - m_\alpha E_\alpha) + 4\sum_{\alpha\beta(\beta'\beta'')}\vartheta_2\begin{pmatrix}\alpha & \beta\\ \beta' & \beta''\end{pmatrix} \times[m_\alpha m_\beta E_\alpha - m_{\beta'}m_{\beta''}\bar{E}_\alpha]\vartheta_2^*\begin{pmatrix}\alpha & \beta\\ \beta' & \beta''\end{pmatrix}, \quad (47)$$

where $\bar{E}_\alpha$ is the average energy of all the states in the band α. $\bar{E}_\alpha$ arises formally if in (46) the index κk is summed over all values, and the factor m_α on the right hand side is dropped. The Hermitian character of ϑ_2, i.e.,

$$\vartheta_2^*(\beta'\beta'',\alpha\beta) = \vartheta_2(\alpha\beta,\beta'\beta''), \quad (48)$$

has been assumed in writing (47). The correlation term in (47) is positive, because $E_\alpha > \bar{E}_\alpha$, i.e., the holes in Ψ_0 have chosen to occupy the highest states available.

In treating the twofold degenerate band, the total density of valence holes is denoted by m, the two orbits by a and b, the spin directions by attaching an arrow to the orbital label. In the ferromagnetic case each band is occupied by $m/2$ holes per lattice site, which gives an average energy E_f according to (46). In the nonferromagnetic case the corresponding quantities are $m/4$ and E_n. In computing the Coulomb-interaction terms according to (43) we shall assume that only the ordinary Coulomb repulsion and exchange integrals are different from zero. Therefore, only the terms containing

$$\begin{aligned} A &= e^2\int\int\frac{dxdy}{|x-y|}|\varphi_a(x)|^2|\varphi_a(y)|^2,\\ B &= e^2\int\int\frac{dxdy}{|x-y|}|\varphi_b(x)|^2|\varphi_b(y)|^2,\\ C &= e^2\int\int\frac{dxdy}{|x-y|}|\varphi_a(x)|^2|\varphi_b(y)|^2,\\ J &= e^2\int\int\frac{dxdy}{|x-y|}\varphi_a(x)\varphi^*{}_b(x)\cdot\varphi_b(y)\varphi^*{}_a(y), \end{aligned} \quad (49)$$

are taken into account for the evaluation of (43). Correspondingly, among the ϑ_2 only those will be assumed to be different from zero which immediately relate to the above integrals. Therefore, the following abbreviations are introduced for the nonferromagnetic case

$$\begin{aligned} 2\vartheta_2\begin{pmatrix}a\uparrow & a\downarrow\\ a\uparrow & a\downarrow\end{pmatrix} &= 1-\epsilon_A,\\ 2\vartheta_2\begin{pmatrix}b\uparrow & b\downarrow\\ b\uparrow & b\downarrow\end{pmatrix} &= 1-\epsilon_B,\\ 2\vartheta_2\begin{pmatrix}a\uparrow & b\downarrow\\ a\uparrow & b\downarrow\end{pmatrix} &= 2\vartheta_2\begin{pmatrix}a\downarrow & b\uparrow\\ a\downarrow & b\uparrow\end{pmatrix} = 1-\epsilon_C,\\ 2\vartheta_2\begin{pmatrix}a\uparrow & b\uparrow\\ a\uparrow & b\uparrow\end{pmatrix} &= 2\vartheta_2\begin{pmatrix}a\downarrow & b\downarrow\\ a\downarrow & b\downarrow\end{pmatrix} = 1-\epsilon_D,\\ 2\vartheta_2\begin{pmatrix}a\uparrow & b\downarrow\\ a\downarrow & b\uparrow\end{pmatrix} &= 2\vartheta_2\begin{pmatrix}a\downarrow & b\uparrow\\ a\uparrow & b\downarrow\end{pmatrix} = \epsilon_J. \end{aligned} \quad (50)$$

In the ferromagnetic case only

$$2\vartheta_2\begin{pmatrix}a\uparrow & b\uparrow\\ a\uparrow & b\uparrow\end{pmatrix} = 1-\epsilon \text{ occurs.}$$

In this terminology the expectation value for the energy per lattice site is given in the ferromagnetic case by

$$4\bar{E} - mE_f + (m^2/2)(1-\epsilon)^2(E_f-\bar{E}) + (m^2/4)\epsilon^2(C-J), \quad (51)$$

and in the nonferromagnetic case by

$$\begin{aligned} 4\bar{E} - mE_n &+ (m^2/8)(E_n-\bar{E})\{(1-\epsilon_A)^2+(1-\epsilon_B)^2\\ &+2(1-\epsilon_C)^2+2(1-\epsilon_D)^2+2\epsilon_J^2\}\\ &+(m^2/16)\{A\epsilon_A^2+B\epsilon_B^2+2(C-J)\epsilon_D^2\\ &+2C(\epsilon_C^2+\epsilon_J^2)-4J\epsilon_C\epsilon_J\}. \end{aligned} \quad (52)$$

The remaining terms in the Coulomb repulsion are the same in the ferromagnetic and in the nonferromagnetic situations.

The minima are reached for

$$\epsilon = \frac{2(E_f-\bar{E})}{C-J+2(E_f-\bar{E})}, \quad (53)$$

with the energy (51) becoming

$$4\bar{E} - mE_f + \frac{m^2}{4}\frac{2(E_f-\bar{E})(C-J)}{C-J+2(E_f-\bar{E})}. \quad (54)$$

Since E_f is very near the top of the band structure if m is small, the quantity $2(E_f-\bar{E})$ can be interpreted as the width of the band. For the nonferromagnetic case one finds

$$\begin{aligned} \epsilon_A &= \frac{2(E_n-\bar{E})}{A+2(E_n-\bar{E})},\\ \epsilon_B &= \frac{2(E_n-\bar{E})}{B+2(E_n-\bar{E})},\\ \epsilon_C &= \frac{2(E_n-\bar{E})(C+2E_n-2\bar{E})}{(C+2E_n-2\bar{E})^2-J^2},\\ \epsilon_D &= \frac{2(E_n-\bar{E})}{C-J+2(E_n-\bar{E})},\\ \epsilon_J &= \frac{2(E_n-\bar{E})J}{(C+2E_n-\bar{E})^2-J^2} \end{aligned} \quad (55)$$

with the energy (52) equal to

$$\begin{aligned} 4\bar{E} - mE_n + \frac{m^2}{16}\Big\{&\frac{2(E_n-\bar{E})A}{A+2(E_n-\bar{E})}\\ &+\frac{2(E_n-\bar{E})B}{B+2(E_n-\bar{E})}+2\frac{2(E_n-\bar{E})(C-J)}{C-J+2(E_n-\bar{E})}\\ &+2\frac{C(C+2E_n-2\bar{E})-J^2}{(C+2E_n-2\bar{E})^2-J^2}2(E_n-\bar{E})\Big\}. \end{aligned} \quad (56)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The quantities ϵ are related by the equation $\epsilon_D-\epsilon_C=\epsilon_J$. By virtue of this relation the sum of all ϑ_2 operators commutes with every component of the total spin of the atom. A group of atomic states, e.g., the triplets made with the orbitals a and b which differ only in the direction of the total spin, occur with the same weight in Ψ, provided they had the same weight in Ψ_0. The expectation value for any component of the total spin vanishes at each lattice site, in Ψ as well as in Ψ_0. But the expectation value for the square of the total spin at any given lattice site may have changed.

The occurrence of ϵ_J can be viewed in the following way. $\bar{\rho}_2$ as given by (43) becomes diagonal in the two-particles wave functions $(a\uparrow a\downarrow)$, $(b\uparrow b\downarrow)$, $(a\uparrow b\uparrow)$, $(a\downarrow b\downarrow)$, $(a\uparrow b\downarrow)+(a\downarrow b\uparrow)$, $(a\uparrow b\downarrow)-(a\downarrow b\uparrow)$. The coefficients of the last two are $m^2(1-\epsilon_C+\epsilon_J)^2/16$ and $m^2(1-\epsilon_C-\epsilon_J)^2/16$. The correlated wave function Ψ prefers therefore the triplet over the singlet, even if there is no ferromagnetism present. Ψ is an example of how Hund's rule can be enforced even in unfilled bands.

In comparing (54) and (56) one notices that the last term in (54) is still smaller than the last two terms in (56) provided $(E_n-\bar{E})<(E_f-\bar{E})(1+J/C)$. This condition should be satisfied since one expects the d band to have a reasonable energy density at its top, so that $E_n-E_f\ll E_f-\bar{E}$ for small density of holes. If one assumes for simplicity's sake that the last term in (54) is equal to the last two terms in (56), and also that $A=B$, then the question is whether or not $m(E_n-E_f)$ is smaller than $m^2(E_n-\bar{E})A/4(A+2E_n-2\bar{E})$. If so, the ferromagnetic correlated state Ψ_f has a lower expectation value for the total energy than the nonferromagnetic Ψ_n.

If the band is parabolic at its top, $E_n-E_f\sim m^{2/3}$, and ferromagnetism is not possible in the present theory. If the density of states as a function of energy reaches a nonvanishing value ν_0 at the top of the band, $m(E_n-E_f)$, is quadratic in m, and ferromagnetism occurs, provided

$$2\nu_0(E_n-\bar{E})>[1+2(E_n-\bar{E})/A]. \tag{57}$$

The density ν_0 is meant to refer to only one of the four bands, arising from the spin and the orbital degeneracy.

One may consider $2(E_n-\bar{E})$ as the width of the d band; $2\nu_0(E_n-\bar{E})$ is then the density of states at the top measured against the density of states resulting from a uniform distribution of the states throughout the band. The inequality (57) requires ν_0 to be larger than the average of ν throughout the band by a factor $2[1+2(E_n-\bar{E})/A]$. In particular, this condition is satisfied if there is a peak in the density of states at its top, such as one may find in the fcc cubic structure.

It may be interesting to note the following: In a nondegenerate band with the Coulomb repulsion integral between localized states given by A, the condition for ferromagnetism is again (57). However, the ferromagnetic state in the nondegenerate band does not have the additional advantage which is due to the last term in (54) being smaller than the last two terms in (56). This advantage of the degenerate band over the nondegenerate band vanishes only in the case of very strong Coulomb repulsion between the two degenerate orbits. Therefore, the results of this section bring out the importance of degeneracy for achieving ferromagnetism.

ACKNOWLEDGMENTS

The author wishes to express his indebtedness to Professor R. Brout, Dr. S. Nettel, and Dr. H. Thomas for many very helpful discussions and suggestions.

APPENDIX A

Equation (29) has to be solved in the special case where $O^{(\nu)}$ differs from zero only if the set of band indices $(\gamma_1\cdots\gamma_\nu)$ coincides with the set $(\alpha_1\cdots\alpha_\nu)$, the factor associated with $O^{(\nu)}$ being independent of the set $(\gamma_1\cdots\gamma_\nu)=(\alpha_1\cdots\alpha_\nu)$. The ϑ_μ will be assumed to have these same properties. It is then not necessary to write down two sets of indices for every operator $O^{(\nu)}$ or ϑ_μ, and even the common set of band indices no longer has to be mentioned since the factors associated with $O^{(\nu)}$ depend only on the number ν for $O^{(\nu)}$ or μ for ϑ_μ. We shall, therefore, only discuss these factors. Since the indices $(\gamma_1\cdots\gamma_\mu)$ in ϑ_μ can still be permuted with respect to the indices $(\alpha_1\cdots\alpha_\mu)$ which are always supposed to be in some standard order; the quantity $\mu!$ ϑ_μ will always occur in the forthcoming argument.

The formulas (28) and (29) have now to be matched. We take a set of m indices $\beta_1,\cdots,\beta_m$ and collect the coefficients of $a^\dagger_{\beta_m}\cdots a^\dagger_{\beta_1}b_{\beta_1}\cdots b_{\beta_m}$ in (29) and in the sum over ν of (28), where we have to insert the definition (19) of the projection operator P_β. In this manner we find that

$$\frac{1}{\omega}\sum_{\substack{\mu_1+2\mu_2+\cdots=m,\\ \mu_1\geq 0,\mu_2\geq 0,\cdots}}\frac{m!}{\mu_1!\mu_2!\cdots}\vartheta_1^{\mu_1}\vartheta_2^{\mu_2}\cdots = \sum_{\nu=0}^{m}(-1)^{m-\nu}\binom{m}{\nu}O^{(\nu)}. \tag{A1}$$

In order to solve this system of equations we multiply each side with $z^m/m!$ where z is an arbitrary real number, and add over all values of m. This gives

$$\frac{1}{\omega}\exp(\vartheta_1 z+\vartheta_2 z^2+\cdots)=e^{-z}\sum_{\nu=0}^{\infty}\frac{z^\nu}{\nu!}O^{(\nu)}, \tag{A2}$$

after some elementary algebra. The terms with $m=0$ were put equal to one; ω can be fixed by the condition $O^{(0)}=1/\omega$.

The case $O^{(0)}=0$ presents difficulties in this formulation. Although they could have been avoided, it seems important to bring out quite clearly that equation (29) is useless as $O^{(0)}$ tends to zero. $O^{(0)}=0$ means that any

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

configuration with at least one lattice site completely empty is excluded in the correlated wave function. A requirement like $O^{(0)}=0$ is absolutely reasonable when there are at least as many particles as lattice sites. The present theory can, therefore, be expected to break down if it is physically necessary to discriminate against configurations with completely empty lattice sites. However, as long as such a condition does not have to be imposed, we can set $\omega=1$, since only the relative weights between the $O^{(\nu)}$ and between the ϑ_μ are physically significant.

The case of greatest interest is

$$O^{(0)}=O^{(1)}=1, \quad O^{(2)}=\epsilon, \tag{A3}$$

from which we obtain

$$\vartheta_1 z+\vartheta_2 z^2+\cdots=\ln\left(1+z+\epsilon\frac{z^2}{2}\right)-z$$

$$=\sum_{n=1}^{\infty}(-1)^{n-1}z^n\sum_{\nu=0}^{\nu=n-1}\frac{1}{n-\nu}\binom{n-\nu}{\nu}\left(\frac{\epsilon}{2}\right)^{\nu}-z. \tag{A4}$$

This gives for $\vartheta_1, \vartheta_2, \cdots$ the values

$$\vartheta_1=0, \qquad \vartheta_2=-\tfrac{1}{2}(1-\epsilon),$$
$$\vartheta_3=\tfrac{1}{3}-(\epsilon/2), \quad \vartheta_4=-\tfrac{1}{4}(1-6\epsilon+\epsilon^2), \quad \text{etc.} \tag{A5}$$

The chain of equations can be broken off at any point.

If ω is left explicitly in the relation between ϑ_μ and $O^{(\nu)}$ the conditions

$$O^{(0)}=\frac{1}{\omega}, \quad O^{(1)}=1, \quad O^{(2)}=\epsilon, \tag{A6}$$

lead to the equations

$$\ln\left(1+\omega z+\epsilon\omega\frac{z^2}{2}\right)-z=\vartheta_1 z+\vartheta_2 z^2+\cdots \tag{A7}$$

$$\vartheta_1=\omega-1, \qquad \vartheta_2=-\frac{\omega}{2}(\omega-\epsilon),$$

$$\vartheta_3=\omega^2\left(\frac{\omega}{3}-\frac{\epsilon}{2}\right), \quad \vartheta_4=-\frac{\omega^2}{4}(\omega^2-6\omega\epsilon+\epsilon^2), \quad \text{etc.} \tag{A8}$$

The case $\epsilon=1$, $\omega\to\infty$ is physically sound, but mathematically useless in terms of ϑ_1, ϑ_2, etc.

APPENDIX B

The two diagrams of Fig. 3 which contribute to (37) can be interpreted as follows. The first diagram gives ρ_1 for the wave function where all Bloch states κk are filled. The second diagram subtracts from the first the contribution

$$\sum_{(\kappa k)}\psi_{\kappa k}(x)\psi^*_{\kappa k}(y) \tag{B1}$$

which comes from the empty Bloch states of Ψ_0.

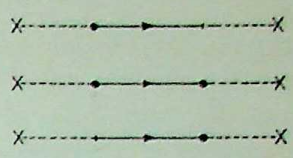


FIG. 9. Simple ρ_1 diagrams containing ϑ_1.

If we combine this last contribution to ρ_1 with the three contributions of Fig. 9, we can write the result as

$$\sum_{(\kappa k)}\bar{\psi}_{\kappa k}(x)\bar{\psi}^*_{\kappa k}(y), \tag{B2}$$

with the new Bloch state

$$\psi_{\lambda k}(x)=\psi_{\kappa k}(x)+\sum_{\alpha\beta\gamma\lambda}\psi_{\lambda k}(x)u_{\alpha\lambda}(k)S_{\alpha\beta k}\vartheta_1(\beta\gamma)u_{\gamma\kappa}(k). \tag{B3}$$

We have used the Fourier transform

$$S_{\alpha\beta k}=\sum_{g}S(\alpha 0,\beta g)e^{ikg}. \tag{B4}$$

The result of all the other diagrams with only ϑ_1 can be stated with the help of a matrix $T_{\kappa\lambda}(k)$ which has for every k as many rows and columns as there are holes with wave vector k in Ψ_0. We have

$$T_{\kappa\lambda}=\delta_{\kappa\lambda}+\sum_{\alpha\beta\gamma}\{u_{\alpha\kappa}S_{\alpha\beta}\vartheta_1(\beta\gamma)u^*_{\gamma\lambda}+u_{\alpha\kappa}\vartheta^*_1(\beta\alpha)S_{\beta\gamma}u^*_{\gamma\lambda}$$
$$+u_{\alpha\kappa}\vartheta^*_1(\beta'\alpha)S_{\beta'\beta''}\vartheta_1(\beta''\gamma)u^*_{\gamma\lambda}\}, \tag{B5}$$

where all the sums over band indices are taken at fixed wave vector k. Instead of subtracting (B1) or (B2) from the ρ_1 of a wave function with all Bloch states κk filled, one has now to subtract

$$\sum_{\kappa\lambda k}\bar{\psi}_{\kappa k}(x)T^{-1}_{\kappa\lambda}\bar{\psi}^*_{\lambda k}(y). \tag{B6}$$

The inverse of T which is inserted in (B6) as compared to (B2), simply normalizes the functions $\bar{\psi}_{\kappa k}(x)$. The term coming from the second diagram of Fig. 3 has, therefore, been replaced by a very similar term, namely one that arises from replacing the Bloch states $\psi_{\kappa k}(x)$ by the properly orthonormalized modification of $\bar{\psi}_{\kappa k}(x)$. Indeed, the states $\bar{\psi}_{\kappa k}(x)$ are not orthonormal, but one can find linear combinations which are orthonormal, say $\tilde{\psi}_{\kappa k}(x)$, and (B6) transforms then to

$$\sum_{(\kappa k)}\tilde{\psi}_{\kappa k}(x)\tilde{\psi}^*_{\kappa k}(y). \tag{B7}$$

Let us now examine an arbitrary diagram without any ϑ_1 in it. The particle lines connecting two vertices

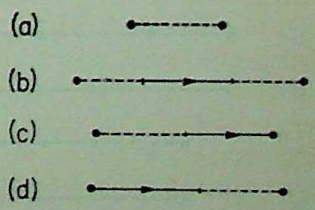


FIG. 10. The four possible particle lines between vertices if $\vartheta_1=0$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

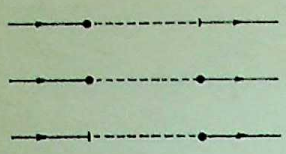


FIG. 11. The three possible insertions of ϑ_1 into a solid line.

or connecting a vertex to an external point are of the four different types shown in Fig. 10. (An external point is equivalent to a connection with a vertex.) The occurrences of ϑ_1 can arise in the following two ways.

First, the solid lines in Figs. 10(b), 10(c), and 10(d) are broken up an arbitrary number of times by inserting one of the three pieces shown in Fig. 11. The effect of these insertions is equivalent to the substitution

$$W(\alpha f,\gamma h)=\frac{1}{L}\sum_{(\kappa k)} U^*_{\alpha f,\kappa k}U_{\gamma h,\kappa k}\rightarrow \frac{1}{L}\sum_{(\kappa\lambda k)} U^*_{\alpha f,\kappa k}T^{-1}{}_{\kappa\lambda}U_{\gamma h,\lambda k}. \quad \text{(B8)}$$

Second, a ϑ_1 can be inserted at the ends of the solid lines in Figs. 10(b), 10(c), and 10(d) without breaking them. This gives the five diagrams listed in Fig. 12.

Combining Fig. 10(b) with Figs. 12(a), 12(b), and 12(c) is tantamount to substituting

$$W(G''G')\rightarrow \sum_{F''F'}[\delta_{F''G'}+\vartheta^*{}_1(F''G')]\times W(F''F')[\delta_{F'G'}+\vartheta_1(F'G')] \quad \text{(B9)}$$

in the diagram 10(b). The substitution (B9) in Fig. 10(c) or Fig. 10(d) can be justified only if the interpretation of the vertex to the right of Figs. 10(c), 10(d), 12(d), and 12(e) is changed by introducing a fictitious label G'' into Figs. 10(c) and 10(d), and G' into Figs. 12(d) and 12(e). This is done by the substitutions

$$\vartheta_2(H'_1H'_2,F'_1F'_2)\rightarrow \sum_{G'_1G'_2}\tilde{\vartheta}_2(H'_1H'_2,G'_1G'_2)\times[\delta_{G'_1F'_1}+\vartheta_1(G'_1F'_1)][\delta_{G'_2F'_2}+\vartheta_1(G'_2F'_2)],$$

$$\vartheta^*{}_2(H''_1H''_2,F''_1F''_2)\rightarrow \sum_{G''_1G''_2}[\delta_{F''_1G''_1}+\vartheta_1(F''_1G''_1)]\times[\delta_{F''_2G''_2}+\vartheta_1(F''_2G''_2)]\tilde{\vartheta}^*{}_2(H''_1H''_2,G''_1G''_2), \quad \text{(B10)}$$

and similar formulas for ϑ_3, etc. If the vertices in Figs. 10(c) and 10(d) are interpreted as $\tilde{\vartheta}_2$, $\tilde{\vartheta}_3$, etc., then the combined effect of Figs. 10(c) and 10(d) as well as 12(d) and 12(e) is again described by the substitution (B9). In all three diagrams 10(b), 10(c), and 10(d), the effect of ϑ_1 insertions is now described by the combined substitutions (B8) and (B9).

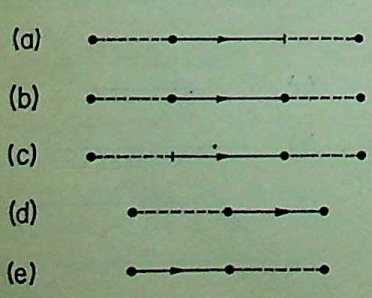


FIG. 12. The five possible ϑ_1 insertions without breaking a solid line.

Now, it follows immediately from (4) that

$$\sqrt{L}\int\psi^*{}_K(x)\psi_G(x)dx=U_{GK}, \quad \text{(B11)}$$

which shows with (32) that

$$W(F''F')=\int\psi^*{}_{F''}(x)dx\times\int\psi_{F'}(y)dy\sum_{(K)}\psi_K(x)\psi^*{}_K(y). \quad \text{(B12)}$$

K is used as abbreviation for the index pair κk. On the other hand, if we insert into this formula the sum over all K, rather than only the K of the holes, then we can write

$$\begin{aligned}\frac{1}{L}\sum_K U^*{}_{F''K}U_{F'K}&=\int\psi^*{}_{F''}(x)dx\int\psi_{F'}(y)dy\times\sum_{G'G''}\psi_{G'}(x)S(G'G'')\psi^*{}_{G''}(y)\\&=\sum_{G'G''}R(F''G')S(G'G'')R(G''G')\\&=R(F''F'). \quad \text{(B13)}\end{aligned}$$

If we take the Fourier transform as in (B4) and insert (40), it follows that

$$\sum_\kappa u^*{}_{\alpha\kappa}(k)u_{\gamma\kappa}(k)=R_{\alpha\gamma k}=\sum_g R(\alpha o,\gamma g)e^{ikg}. \quad \text{(B14)}$$

From this last relation, we find that

$$\sqrt{L}\int\bar{\psi}^*{}_K(x)\psi_G(x)dx=U_{GK}+\sum_F\vartheta_1(GF)U_{FK}. \quad \text{(B15)}$$

Therefore, the combination of (B8) and (B9), i.e., the elimination of all ϑ_1 from the diagrams, is effected by substituting

$$W(F''F')\rightarrow\int\psi^*{}_{F''}(x)dx\int\psi_{F'}(y)dy\times\sum_{(\kappa\lambda k)}\bar{\psi}_{\kappa k}(x)T^{-1}{}_{\kappa\lambda}\bar{\psi}^*{}_{\lambda k}(y) \quad \text{(B16)}$$

and replacing all $\vartheta_2, \vartheta_3, \cdots$ by $\tilde{\vartheta}_2, \tilde{\vartheta}_3, \cdots$.

APPENDIX C

The Bloch waves $\psi_{\kappa k}(x)$ are assumed to arise from some secular problem as a linear combination of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

functions $\psi_{\beta g}(x)$. Under a coordinate transformation $x=Tx'+t$ compatible with the lattice, one has

$$\psi_{\kappa k}(x)=e^{ikt}\psi_{\kappa k'}(x') \quad \text{with} \quad k=Tk'. \tag{C1}$$

Also one has for the localized orbits

$$\psi_{\beta g}(x)=\sum_{\alpha} T_{\beta\alpha}\psi_{\alpha f} \quad \text{with} \quad g=Tf+t, \tag{C2}$$

where the matrix $T_{\beta\alpha}$ gives the transformation of the orbits corresponding to the coordinate transformation T.

In view of the orthonormality of the functions $\psi_{\kappa k}(x)$ one finds for the coefficients U_{GK} of (4) the formula

$$\frac{1}{\sqrt{L}}U_{GK}=\int \psi^*{}_K(x)\psi_G(x)dx. \tag{C3}$$

If one inserts (C1) and (C2) it follows that

$$U_{GK}=e^{-ikt}\sum_{\alpha} T_{\beta\alpha}U_{FK'}, \tag{C4}$$

with $F=(\alpha f)$, $G=(\beta g)$, $K=(\kappa k)$, $K'=(\kappa k')$. Any expression depending on $U_{GK}, \psi_K, \psi_G, K, G$, etc., can be written in terms of $U_{FK'}, \psi_{K'}, \psi_F, K', F$, etc., with the help of (C1), (C2) and (C4). In the new expression in which only the latter quantities and the coefficients $T_{\beta\alpha}$ occur, it is possible to replace again $F\to G$, $K'\to K$, etc. If the resulting expression coincides with the original one, it is called invariant with respect to the transformation T.

Now one has for W as defined in (32) the transformation

$$W(G''G')=\sum_{\alpha'\alpha''} T^*{}_{\beta''\alpha''}T_{\beta'\alpha'}W(F)'F'' \tag{C5}$$

according to (C4). In order to obtain (39) we now put $t=0$, $g''=g'=f''=f'$. The arguments in W are then simply $\beta''\beta'$ and $\alpha''\alpha'$.

The formula (C5) is now averaged over all symmetry operations of the lattice which leave the lattice site $g=f$ fixed. This gives

$$W(\beta''\beta')=\sum_{\alpha'\alpha''}\left(\frac{1}{l}\sum_{T} T^*{}_{\beta''\alpha''}T_{\beta'\alpha'}\right)W(\alpha''\alpha'). \tag{C6}$$

If the orbits have been grouped according to the irreducible representation to which they belong, it is possible to apply one of the important results of group representation theory.[7]

$$\frac{1}{l}\sum_{T} D^{(j'')}{}_{\beta''\alpha''}(T)D^{(j')}{}_{\beta'\alpha'}(T)=\frac{1}{j}\delta_{j'j''}\delta_{\beta'\beta''}\delta_{\alpha'\alpha''}, \tag{C7}$$

where $D^{(j)}{}_{\beta\alpha}(T)$ is the matrix representing T in the representation j, the indices α', β', α'', β'' being restricted to the values relevant to j' and j''. j is the dimension of the representation j. l is the number of operations in the group. If (C7) is inserted into (C6) the formula (39) and the explanation preceding it follow immediately.

[7] Cf. Eugene P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, translated by T. T. Griffin (Academic Press, Inc., New York, 1959), p. 83.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Mechanism of the Ferrimagnetic to Antiferromagnetic Transition in $Mn_{2-x}Cr_xSb$

H. S. JARRETT

E. I. du Pont de Nemours and Company, Central Research Department, Wilmington, Delaware*

(Received 16 December 1963)

The mechanism of the first-order ferrimagnetic to antiferromagnetic transition (exchange inversion) in $Mn_{2-x}Cr_xSb$, $x<0.41$, was investigated by exchange magnetostriction. A second-order term has been included in the strain energy which gives an effective elastic constant dependent on magnetization and magnetic ordering. It also introduces a discontinuous change in the magnetization at the transition and hence a contribution to the entropy change in addition to that caused by the change of lattice dimension. For $Mn_{2-x}Cr_xSb$, the effect of this term is small but necessary to obtain the qualitative behavior of the discontinuous change in lattice dimension and the shift in the Curie temperature with composition. The anisotropy behavior of $Mn_{2-x}Cr_xSb$ was investigated for various values of x, and it is shown that an anisotropy in the critical lattice dimension of only 4×10^{-4} Å accounts for the observed change in magnetic anisotropy with exchange inversion that occurs for transition temperatures between 200 and 350°K. For transition temperatures above 350°K, anisotropy is positive, and spin-flopping measurements can be made on the antiferromagnetic state. An expression for the critical field in the presence of exchange inversion is obtained and compared with experiment.

INTRODUCTION

THE basic features of exchange inversion have been investigated theoretically by Kittel.[1] In this paper Kittel's theory is extended to include Curie-point shifts, magnetization, entropy, and magnetic anisotropy and is applied to Cr-modified Mn_2Sb. Kittel proposes that a transformation from ferromagnetic (F) to antiferromagnetic (AF) spin ordering will take place by a first-order process when there is, during normal thermal contraction, a critical lattice dimension c_c at which the sign of the exchange interaction changes from F to AF. Accompanying the first-order transition is a discontinuous change in dimension of the crystalline unit cell. The magnitude of the discontinuous change, as Kittel has pointed out, is equal to the difference between the magnetoelastic expansion of the ferromagnetically ordered lattice and the contraction of the antiferromagnetically ordered lattice. Experimental data that have been obtained support qualitatively the essential features of this theory.

The ternary composition $Mn_{2-x}Cr_xSb$ is an ideal magnetic material on which to check semiquantitatively the features of an exchange-inversion theory. Mn_2Sb forms sets of 3-layer sheets[2] of strongly coupled spins ($J/k\sim450$°K); the spins between sets are only weakly coupled by exchange forces ($J/k\sim80$°K at the Curie temperature). The net saturation moment of this 3-layer set is approximately 0.9 Bohr magneton.[3] We shall assume that detailed knowledge of the interatomic exchange forces is not needed and that this 3-layer set may be replaced with a single plane of spins strongly coupled ferromagnetically. Effects of magnetoelastic energy within this plane will also be neglected.

Thus, the model representing Mn_2Sb is a stack of weakly coupled ferromagnetic planes, the interplanar exchange changing sign at some accessible lattice dimension. A main function of the Cr modifier is to provide an initial lattice contraction so that, depending on the amount present, the critical lattice dimension for exchange inversion occurs at some temperature below the Curie temperature.

FREE ENERGY

In order to account for the exchange energy dependence on lattice dimension, the exchange energy is expanded in a Taylor series to quadratic terms about the critical lattice dimension c_c at which the exchange energy vanishes:

$$J(c)=\rho(c-c_c)+(\rho'/2)(c-c_c)^2. \quad (1)$$

Here ρ and ρ' are the first and second derivatives of the exchange interaction evaluated at c_c. This expansion for the exchange energy is identical to that used by Kittel except that he considered only the linear dependence of J on lattice parameter.

When the elastic energy $\frac{1}{2}R(c-c_T)^2$ is included, the interplanar free energy per unit volume becomes

$$F/V=\tfrac{1}{2}R(c-c_T)^2 -[\rho(c-c_c)+(\rho'/2)(c-c_c)^2]m_1m_2\cos\phi. \quad (2)$$

Here, R represents the elastic constant divided by the square of the average lattice dimension, and c_T is the lattice dimension in the absence of magnetic forces. The parameter m_i is the reduced sublattice magnetization obtained by normalizing the magnetization to the saturation value M_0 at absolute zero.

The equilibrium value of c is given by $\partial F/\partial c=0$, which yields

$$c=c_T+\left[1-\frac{\rho' m_1m_2}{R}\cos\phi\right]^{-1} \times\left[\frac{\rho m_1m_2}{R}\cos\phi+(c_T-c_c)\frac{\rho' m_1m_2}{R}\cos\phi\right] \quad (3)$$

* Contribution No. 830.

[1] C. Kittel, Phys. Rev. 120, 335 (1960).

[2] L. Heaton and N. S. Gingrich, Acta Cryst. 8, 207 (1955).

[3] C. Guillaud, thesis, University of Strasbourg, 1943 (unpublished). L. Néel, Ann. Phys. 3, 137 (1948).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and, upon substitution into the free energy, Eq. (2), we obtain

$$\frac{F}{VR}=\left\{-\frac{1}{2}\left(\frac{\rho m_1 m_2}{R}\right)^2\cos^2\phi-\frac{\rho}{R}m_1m_2(c_T-c_c)\right.$$
$$\left.\times\left[1+\frac{\rho'}{2\rho}(c_T-c_c)\right]\cos\phi\right\}\left[1-\frac{\rho' m_1 m_2}{R}\cos\phi\right]^{-1}. \quad (4)$$

Both the elastic energy and the second term in the exchange energy appear quadratically. A quadratic term in the exchange produces an effective elastic constant that is dependent on the spin orientation.

The condition for a transition to occur between F and AF spin configurations is no longer $c_T=c_c$ of the Kittel theory but is readily obtained by equating the free energy for F ordering to the free energy for AF ordering to give

$$c_T-c_c=\frac{\rho}{\rho'}\left\{-1+\left[1-\left(\frac{\rho' m_1 m_2}{R}\right)^2\right]^{1/2}\right\}. \quad (5)$$

This expression reduces to the Kittel condition when $\rho'=0$, and approaches the Kittel condition as the exchange-inversion temperature approaches the Curie point, where the sublattice magnetization is small. We now obtain from Eq. (3) the difference in interplanar spacing Δc between the F and AF states,

$$\Delta c=c^{\mathrm{F}}-c^{\mathrm{AF}}=\frac{2\rho m_1 m_2}{R[1-(\rho' m_1 m_2/R)^2]}\left[1+\frac{\rho'}{\rho}(c_T-c_c)\right]. \quad (6)$$

Substitution of Eq. (5) for the transition condition gives the discontinuous change in lattice dimension at the transition

$$\Delta c_s=\frac{2\rho m_1 m_2}{R}\left[1-\left(\frac{\rho' m_1 m_2}{R}\right)^2\right]^{-1/2}, \quad (7)$$

which is similar to the Kittel theory but is modified by the quadratic exchange term.

So far, this development has paralleled that of Kittel, but the results given demonstrate how the quadratic exchange term alters related conditions of the linear exchange theory. These expressions will also be necessary when attempting a fit to the experimental data using ρ/R and ρ'/R as adjustable parameters.

The free-energy expression, Eq. (4), is not complete, since the intraplanar exchange energy Jm^2 and magnetic ordering entropy S have not been taken into account. Therefore, the following terms must be added to Eq. (4):

$$-(J_1/2)m_1^2-(J_2/2)m_2^2-kTS_1-kTS_2, \quad (8)$$

where the classical entropy for N spins[4] with reduced moment m is given by

$$\sigma=-\frac{N}{2}\left[\left(\frac{1+m}{2}\right)\ln\left(\frac{1+m}{2}\right)+\left(\frac{1-m}{2}\right)\ln\left(\frac{1-m}{2}\right)\right]$$

for each sublattice.

ENTROPY CHANGE

Addition of the terms (7) to Eq. (4) gives the total free-energy expression of the system including, in a rather disguised form, the lattice entropy. This fact may be demonstrated by differentiating the free energy with respect to temperature and taking the difference between the total entropy for F ordering and AF ordering. At the transition, given by the condition of Eq. (5), we obtain

$$\Delta S=-R\Delta c\left[\frac{\rho'}{R}\frac{1}{m}\frac{\partial m}{\partial T}\frac{\Delta c}{2}+\frac{\partial c_T}{\partial T}\right]. \quad (9)$$

Thus, the change in entropy of the lattice at the transition is included in the dimension c_T. The first term in the bracket of Eq. (9) is of interest as it represents the magnetic contribution of the effective elastic energy to the change in entropy. Magnetization terms do not enter otherwise in Eq. (9) because the sublattice magnetization is assumed to be unchanged at the transition.[5] The magnitude of the first term of Eq. (9) at room temperature is of the order

$$\frac{1}{2}\frac{\rho'}{R}\frac{1}{m}\frac{\partial m}{\partial T}\frac{\Delta c}{c}\approx\frac{0.3}{2}(0.0015)\left(\frac{0.03}{6.5}\right)\approx 10^{-6}/^{\circ}\mathrm{K}$$

for Mn_2Sb, while

$$\frac{1}{c}\frac{\partial c_T}{\partial T}\approx 20\times 10^{-6}/^{\circ}\mathrm{K},$$

and hence the first term is negligible. Therefore, the change in entropy at the F/AF phase transition arises largely from the lattice distortion and becomes

$$\Delta S\simeq -C\left(\frac{\Delta c}{c}\right)\left(\frac{1}{c}\frac{\partial c_T}{\partial T}\right), \quad (10)$$

where C is the appropriate elastic constant in the direction of the strain $\Delta c/c$.

MAGNETIZATION

Minimization of the free energy, Eq. (4) plus (8), with respect to the sublattice magnetization and rearrangement of terms gives

$$m=\tanh\left\{\frac{mM_0}{kT}H_w\right\}, \quad (11)$$

[4] J. S. Smart, Phys. Rev. 90, 55 (1953).

[5] A. E. Austin, E. Adelson, and W. H. Cloud, Phys. Rev. 131, 1511 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$H_w=\frac{2R}{NM_0}\left(\frac{J}{R}+\left\{\left(\frac{\rho m}{R}\right)^2\left(1-\frac{\rho' m^2}{2R}\cos\phi\right)+(c_T-c_c)\left[\rho+\frac{\rho'}{2}(c_T-c_c)\right]\cos\phi\right\}\times\left[1-\frac{\rho' m^2}{R}\cos\phi\right]^{-2}\right). \quad (12)$$

It has been assumed in Eq. (12) that the two sublattices are equivalent and that $J_1=J_2$. The magnetization given by Eq. (11) is actually the Brillouin function for spin $\frac{1}{2}$. This form arises because the entropy expression assumes only two spin orientations. The first term of Eq. (12) represents the Weiss field due to intraplanar exchange, while the second term is due to interplanar exchange and, it is important to note, is dependent on the spin configuration through the $\cos\phi$ term. If the ρ' term were not included, the $\cos\phi$ dependence disappears, and there is no difference in the Weiss fields for F and AF ordering at the transition.

Thus, the magnetization is discontinuous at the transition because of the second-order ρ' term. However, even for fairly large values of $\rho'/R\sim\frac{1}{2}$, the discontinuous change of sublattice magnetization is of the order of a few percent and would be difficult to detect by neutron diffraction, for example. A large change in sublattice magnetization would alter all previously derived expressions, but for practical purposes we may ignore the small change in magnetization as long as $\rho'/R<\frac{1}{2}$. This restriction is no stronger than the initial assumption that the exchange energy dependence on lattice dimension can be represented by Eq. (1).

CURIE TEMPERATURE

Also contained in Eq. (11) is the unusual effect of a temperature-dependent Weiss field. In most molecular field treatments it is assumed that exchange interactions are constant over the entire temperature range below the Curie temperature. In exchange inversion, however, the exchange interaction is dependent on temperature through lattice contraction. Such effects should, therefore, influence the Curie temperature.

The Curie temperature for the magnetization expressed by Eq. (11) is

$$T_c=\frac{R}{Nk}\left\{(c_{T_c}-c_c)\left[\frac{\rho}{R}+\frac{\rho'}{2R}(c_{T_c}-c_c)\right]\cos\phi\right\}+\frac{J}{Nk}, \quad (13)$$

where $(c_{T_c}-c_c)$ indicates that (c_T-c_c) is to be evaluated at T_c. We thus have a self-consistent equation in which T_c is contained explicitly on the left-hand side and implicitly on the right-hand side of Eq. (13). The second term J/Nk in Eq. (13) represents the Curie temperature of the ferrimagnetically ordered 3-layer set of planes, while the first term is the interplanar Curie temperature. The Curie temperature is dependent on $\cos\phi$. For positive ρ the Curie temperature for F ordering is higher than the Curie temperature for AF ordering by twice the interplanar exchange energy. Thus, as temperature is decreased, ferromagnetic ordering first occurs and is the most stable configuration.

APPLICATION TO Mn_2Sb

In this section, the foregoing model is applied to Cr-modified Mn_2Sb.[6] It should be emphasized that the foregoing expressions represent a single composition whose F/AF transition occurs at a temperature T_s. In applying such a model to $Mn_{2-x}Cr_xSb$ for various values of x, in which T_s varies with chromium content, one must investigate whether the behavior of all compositions can be represented by a model whose parameters are independent of composition, and yet whose thermodynamic quantities exhibit the observed functional dependence on composition.

Compositional changes are introduced into the function (c_T-c_c) by assuming that the temperature dependence of c_T remains the same as that of Mn_2Sb, but that the modifying element (chromium) reduces the absolute value of c_T. Thus, the requirement on the modifying element is that the lattice contraction introduced by it is just the amount required to satisfy Eq. (5) at the transition temperature T_s.

There are probably some deviations from this assumption at high chromium content, but the theory is not quantitative enough to warrant consideration of such deviations. At this point, we must mention a metallurgical property of this material. An increase in the value of x does not represent a proportional decrease in the value of c_T. It has been found that the chromium tends to segregate into an ever-present MnSb Widmanstätten precipitate,[7] and there is not a corresponding increase of chromium content in the Mn_2Sb phase proportional to the amount introduced into the melt. It must therefore be assumed that the transition temperature T_s occurs when Eq. (5) is satisfied and that the chromium in $Mn_{2-x}Cr_xSb$ has introduced just the reduction in c_T required to satisfy Eq. (5) at T_s.

We shall also assume that c_c is constant throughout the entire range of transition temperatures. The value of c_c is critical only when calculating the F and AF lattice parameters for a given transition temperature. Otherwise c_c always appears in the expression (c_T-c_c) and, if this difference can be obtained, c_c needs to be known only approximately. A third assumption is that the magnetization of the sublattice of our model is given by the measured magnetization rather than by Eq. (11). As has already been discussed,[5] the spins of the 3-layer

[6] F. J. Darnell, W. H. Cloud, and H. S. Jarrett, Phys. Rev. 130, 647 (1963).

[7] J. D. Wolf and J. E. Hanlon, J. Appl. Phys. 32, 2584 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sets of atomic planes remain ordered antiparallel and appear to be strongly coupled by exchange forces. We assume, therefore, that the details of the magnetization behavior of Mn I and Mn II ions within this plane do not contribute to exchange inversion except through the temperature dependence of the magnetization of the entire 3-layer set. One would not expect that Eq. (11) should represent the magnetization of Mn_2Sb. On the other hand, the conclusions drawn from Eq. (11) by use of the Weiss field is semiquantitatively correct even though the functional form is not. This functional dependence arose because we have used classical statistics in the magnetic entropy and assumed only a single sublattice.

To obtain the parameters ρ/R and ρ'/R, consider the discontinuous change in the c lattice parameter given by Eq. (7). The elastic constant R does not need to be known until absolute energy comparisons are made. The best fit to Δc_s is given by $\rho/R=0.02$ Å and $\rho'/R=0.4$.

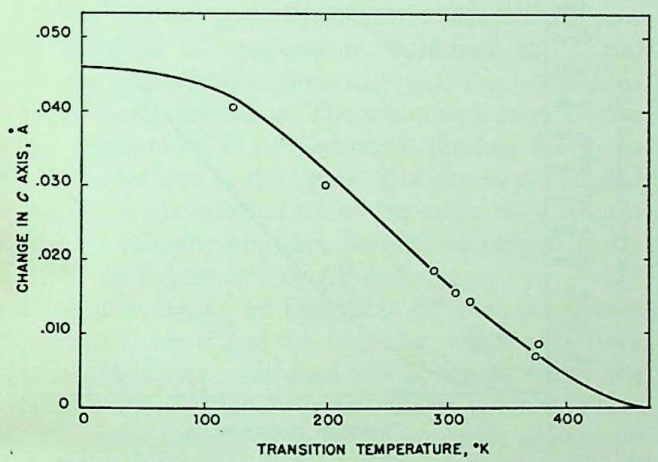


FIG. 1. Discontinuity in c axis as a function of transition temperature T_s. Experimental results (Ref. 6) are shown by the open circles.

Within the limits of experimental error, the value of ρ'/R is not too critical. For example, a 0.1 variation in ρ'/R can be tolerated and still obtain a fit with the experimental data. The value of ρ/R is more critical for only 0.001 Å variation can be tolerated here. This fit of experimental data to the theoretical equation is shown in Fig. 1. Since the introduction of chromium reduces both T_c and the saturation moment, the magnetization of Mn_2Sb cannot be used in Eq. (7). Instead, the magnetization measured at $T=T_s$ of that composition whose F/AF transition temperature is T_s is used. This magnetization is lower than the magnetization of Mn_2Sb at the corresponding temperature.

We may now obtain the temperature dependence of (c_T-c_c) for unmodified Mn_2Sb. From Eq. (3) we obtain c_F, the dimension of the unit cell for ferromagnetic ordering at $\phi=0$. Rearrangement of terms gives

$$(c_T-c_c)=(c_F-c_c)[1-(\rho'/R)m^2]-(\rho/R)m^2, \quad (14)$$

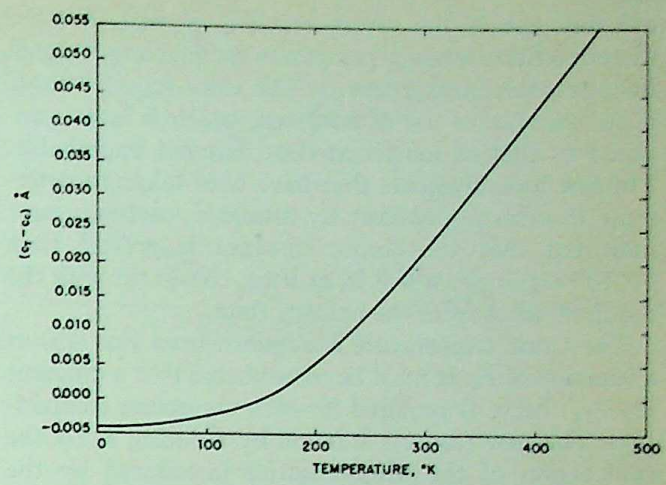


FIG. 2. Dimension (c_T-c_c) as a function of temperature calculated from the thermal contraction of Mn_2Sb.

where the values of the magnetization to be used in Eq. (15) are those of Mn_2Sb. It is seen in Eq. (14) that c_c must be known. The experimentally measured lattice dimension of Mn_2Sb is c_F. The value of c_c within reasonable limits does not affect appreciably the empirically determined temperature dependence of (c_T-c_c). Selection of the value of $c_c=6.506$ Å is discussed later in the paper. Figure 2 shows the temperature dependence of (c_T-c_c) obtained from Eq. (14).

The change in entropy may now be calculated by use of Fig. 2 and Eq. (10). The value of the entropy is dependent only on the rate of change of c_T with temperature. The calculated change in entropy is shown in Fig. 3 in units of the elastic constant. Agreement with the experimentally measured ΔS given by the open circles is good, if it is assumed that the elastic constant C is 4.7×10^{12} ergs/cm³. Some experimental values[8] of ΔS in units of the elastic constant are shown for comparison. Such agreement is also indicative of the validity of the assumption that the temperature dependence of c_T is nearly independent of chromium content. We have been unable to measure an elastic constant along the c

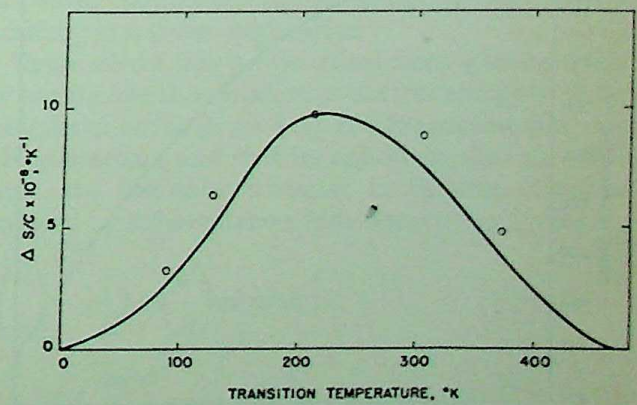


FIG. 3. Entropy change versus transition temperature. The data of Flippen (Ref. 9) are shown for comparison.

[8] R. B. Flippen and F. J. Darnell, J. Appl. Phys. 34, 1094 (1963); W. A. Doerner and R. B. Flippen (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

axis because of the severe cracking normal to the c direction that always accompanies samples large enough for ultrasonic measurements. The value required for C is comparable to C_{11} of tungsten, which is large compared to that of most materials, but not improbable. The few measurements that have been taken to determine the elastic constant by ultrasonic methods have indicated that the elastic constant is greater than 3×10^{12} ergs/cm^3, which is, at least, consistent with the required value from the entropy data.

The Curie temperature is obtained from Eq. (13) as a function of T_s. It must be remembered that a different (c_T-c_c) curve is required for each chromium composition. The new (c_T-c_c) is found by reducing c_T by the contraction of the Mn_2Sb lattice introduced by the chromium. The temperature which satisfies the equality in Eq. (13) is the Curie temperature for the composition. The Curie temperatures for various ternary compositions were thus calculated from the experimentally determined lattice contractions and the results are shown in Fig. 4 plotted against the associated F/AF transition temperature T_s. The measured Curie temperatures are included for comparison.

For the foregoing calculation the intraplanar exchange energy must be known. From the measured Mn_2Sb Curie temperature J/Nk is found to be 470°K by subtracting the value of the first term for Mn_2Sb on the right-hand side of Eq. (13). Actually an unknown parameter is introduced by the scale factor R/Nk, which converts elastic energy to temperature units. Our model compresses the 3-layer set of planes into a single layer. Therefore, the effective density N of the magnetic ions is not known, and the value we select for N influences the value of the intraplanar Curie temperature. However, N cannot be varied widely because the T_c calculated for the modified Mn_2Sb materials would deviate widely from the experimentally measured values. There is no real advantage gained in making a least-squares fit, so the density of magnetic ions in the 3-layer set, 6.2×10^{22}/cm^3, is selected as the density N of ions for the model. The effective value of R is obtained from the change in entropy, Eq. (9), and the scale factor is $R/Nk=1.5\times10^4$ °K/Å^2.

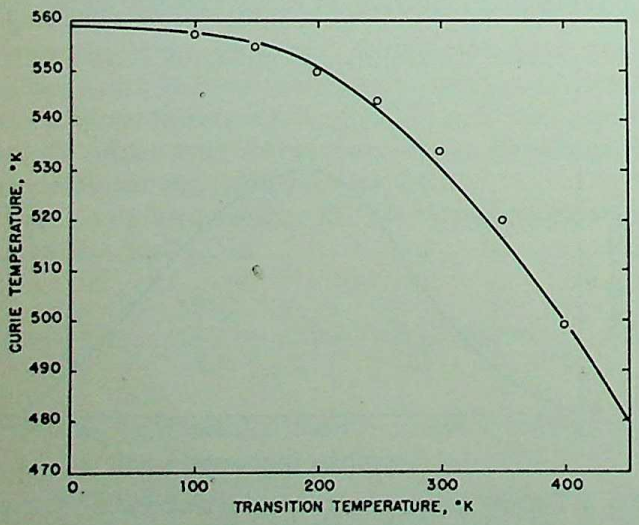


FIG. 4. Dependence of Curie temperature on F/AF transition temperature. Experimental results (Doerner and Flippen, to be submitted) are shown for comparison.

The parameters which have been evaluated allow calculation of the sublattice magnetization from Eq. (11). The sublattice magnetization for (c_T-c_c) given in Fig. 2, which would correspond to unmodified Mn_2Sb, is essentially a Brillouin function with $J=\frac{1}{2}$. For (c_T-c_c) modified to give $T_s>0$°K, the sublattice magnetization differs slightly from the unmodified (c_T-c_c) because of the reduction of Curie temperature, but the shape is still essentially a Brillouin function. The calculated discontinuity in M for $T_s=300$°K is only 0.8%. There is little agreement between the experimental and the theoretical magnetization compared to the rather remarkable agreement in the foregoing calculations. This result is to be expected, of course, since the functional form of the magnetic entropy is based on a molecular field model of a single sublattice.

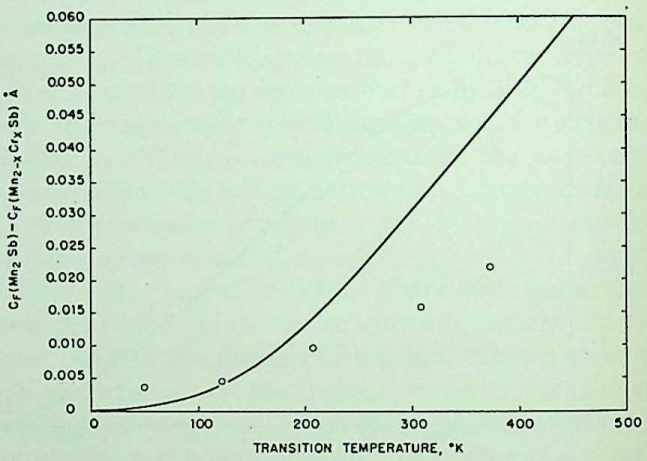


FIG. 5. Reduction of the ferromagnetic lattice dimension with increasing T_s.

A value for c_c has been used in the foregoing calculations, but its derivation has not been established. The expression for c_F is obtained from Eq. (3) for $\phi=0$ and $m=m_s$, the measured value of the magnetization at T_s. The lattice contraction in the F state due to the introduction of chromium is then given by

$$\Delta c_F=\frac{c_T-c_c+(\rho/R)m_s^2}{1-(\rho'/R)m_s^2}+c_c-c_F(Mn_2Sb)$$

and is shown in Fig. 5 with the experimental points obtained from Fig. 1 of Darnell *et al.*[6]

The value of c_c was selected to fit the low T_s materials. It is seen that deviations by a factor of 2 occur for high T_s materials, which indicates that our initial assumptions about the passive role of the chromium may not be entirely valid. Either c_c is not constant over

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the range of transition temperatures or the temperature dependence of c_T changes at high chromium content. However, in view of the success of the theory which assumes the functional form of c_T is not changed with the addition of chromium, it seems more probable that c_c is not constant.

ANISOTROPY

The anisotropy energy has not been taken into account in the foregoing treatment. On the whole, the anisotropy energy is 1 to 2 orders of magnitude smaller than the smallest exchange interaction, and its inclusion has not been necessary. There is however, an effect of anisotropy which is quite apparent. Consider Fig. 10 of Darnell *et al.*[6] It is seen that the temperature T_A at which the anisotropy changes sign for F ordering is not continuous across the T_s boundary into the AF region, but lies on this boundary from 200 to 350°K before entering the AF region. This effect is ascribed to an anisotropy in the interplanar exchange energy. In Fig. 6 the free energy is represented schematically as a linear function of temperature. Although the actual functional dependence is more complex, this representation serves this discussion. The anisotropy contribution is also exaggerated in proportion to the free energy so that it can be seen in the figure. The curves are labeled according to the orientation of the spins with respect to the crystallographic axis, parallel or perpendicular to c, and according to F or AF ordering.

It has been shown by Darnell *et al.*[6] that the anisotropy energies for F and AF ordering, which arise from dipole-dipole and crystal-field effects, are so nearly the same that the small differences can be ignored. Thus, the first-order contribution to the magnetic anisotropy energy to be added to the free energy is the same for F and AF ordering, and anisotropy is continuous across the F/AF transition. However, in second order, the anisotropy mixes with the exchange energy producing an anisotropic exchange. The effect of an anisotropic exchange is shown in Fig. 6 by the difference in the temperature at which $F_{AF}{}^{\perp}$ and $F_F{}^{\perp}$ intersect and $F_{AF}{}^{\parallel}$ and $F_F{}^{\parallel}$ intersect. The free energy at the intersection for parallel ordering is lower than that for perpendicular ordering by the anisotropy energy. For transition temperatures between the $F_{AF}{}^{\perp}$ and $F_F{}^{\perp}$ intersection and the $F_{AF}{}^{\parallel}$ and $F_F{}^{\parallel}$ intersection, Fig. 6(a) results. Since the magnetic system follows the lowest free energy, the F/AF transition takes place with a concomitant transition from positive to negative anisotropy. This situation is the condition for a transition between 200 and 350°K.

Figure 6(b) represents the free-energy relations for F/AF transition temperatures $T_s>350$°K. Here the $F_{AF}{}^{\parallel}$ and $F_F{}^{\perp}$ intersection is at a temperature lower than the F/AF transition temperature T_s, and the anisotropy remains positive at T_s. Figure 6(c) shows the situation for a low F/AF transition temperature,

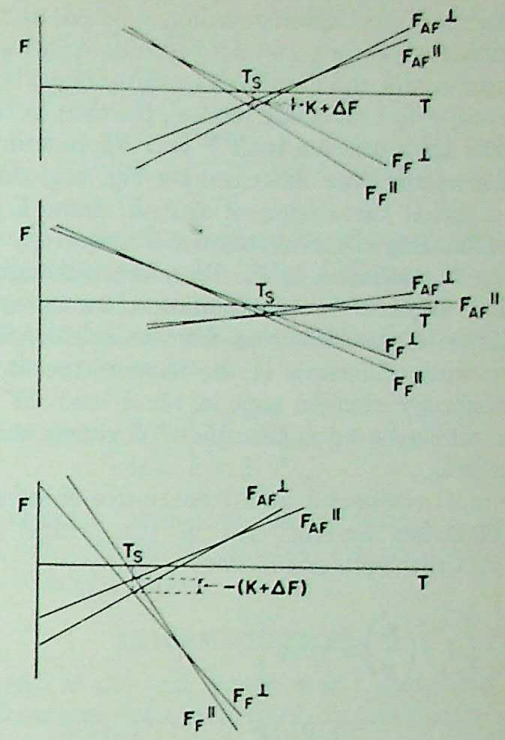


FIG. 6. Schematic representation of the effect of anisotropic exchange on magnetocrystalline anisotropy: (a) 200°K$<T_s<$350°K, (b) $T_s>$350°K, (c) $T_s<$200°K.

$T_s<200$°K. Here the $F_F{}^{\parallel}$ and $F_F{}^{\perp}$ intersection is at a temperature higher than T_s and the anisotropy changes from positive to negative before the F/AF transition occurs.

Relative to the free energy for parallel orientation of spins, the free energy for perpendicular orientation is

$$F_{AF,F}{}^{\perp}=F_{AF,F}{}^{\parallel}+\Delta F_{AF,F}+K^0.$$

The differences between the $F^{\perp}$ and $F^{\parallel}$ curves of Fig. 6 are due to the temperature dependence of the $(\Delta F+K)$ terms. Thus, the anisotropy of Mn_2Sb must contain not only terms of the type described by Darnell *et al.* and designated here by K^0 but also an anisotropic exchange ΔF whose contribution cannot be distinguished experimentally in a given composition.

These effects may be described more quantitatively by considering the variation of the free energy with the interplanar exchange $\rho(c-c_c)$. If it is assumed that only ρ is anisotropic, and that its anisotropy $\Delta\rho=\rho_{\parallel}-\rho_{\perp}$ is small, the first-order deviation in the free energy is obtained by differentiation with respect to ρ, giving

$$\Delta\left(\frac{F}{R}\right)=-\left[1-\left(\frac{\rho'}{R}\right)m^2\cos\phi\right]^{-1}$$
$$\times\left[\left(\frac{\rho m^2}{R}\right)^2\cos^2\phi+\left(\frac{\rho}{R}\right)m^2(c_T-c_c)\cos\phi\right]\frac{\Delta\rho}{\rho}. \quad (15)$$

From Eq. (15), the magnitude of the anisotropy in the free energy ΔF depends on the spin ordering through

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$\cos\phi$. In the Kittel approximation, $\rho'=0$. At the F/AF transition $(c_T-c_c)=0$, and ΔF becomes independent of spin ordering for the F and AF configurations. Without the second-order exchange term ρ', the shift in the free energy is the same for both F and AF ordering, and behavior of the type described by Fig. 6(a) does not arise. Explicit calculation of $\Delta(F/R)$ from Eq. (15) shows that this contribution to the anisotropy of the AF state is smallest near T_s. Therefore, although there may be a second-order contribution to the temperature dependence of the anisotropy from an anisotropic ρ/R, no significant difference in the temperature at which the anisotropy changes sign in the F and AF states occurs. Similarly, an anisotropic ρ'/R suffers the same deficiencies.

The most reasonable additional source of anisotropy is a variation $\Delta c_c=c_c^{\perp}-c_c^{\parallel}$ of the critical lattice dimension with spin orientation,

$$\Delta\left(\frac{F}{R}\right)=\left[1-\left(\frac{\rho'}{R}\right)m^2\cos\phi\right]^{-1}\times\left[\frac{\rho}{R}+\frac{\rho'}{R}(c_T-c_c)\right]m^2\cos\phi\Delta c_c. \quad (16)$$

Here, in the Kittel approximation $\rho'=0$, the incremental changes in the free energy are in opposite direction for the F and AF states, i.e., this contribution to the anisotropy from Eq. (16) changes sign at T_s. Such behavior is just that which is necessary to account for the qualitative arguments summarized by Fig. 6. Thus, the first-order anisotropy constant may be expressed as

$$K=K^0+\rho m^2\Delta c_c\cos\phi, \quad (17)$$

where K^0 contains all other contributions to the anisotropy that are predominately independent of spin ordering, e.g., crystal-field anisotropy and contributions such as expressed by Eq. (15).

In order to show that the additional contribution to the anisotropy in Eq. (17) is plausible, we compare the temperatures at which the anisotropy changes sign for the F and AF states. For the situation described by Fig. 6(a), anisotropy changes sign when exchange inversion occurs for a range in T_s of 150°. By use of the value of ρ/R found previously, Δc_c must be 4×10^{-4} Å. Such a value for Δc_c is reasonable, unfortunately however, it is difficult to detect directly. It is significant, however, that such a small anisotropy in the exchange can play a dominant role in the magnetic anisotropy whenever magnetoelastic effects are present.

SPIN FLOPPING

The appearance of positive anisotropy in the antiferromagnetic state suggests that spin-flopping experiments can be performed on the high-chromium Mn_2Sb compositions. Such experiments have been reported by Flippen.[9] Spin flopping has not been treated in an exchange-inversion material, and we outline briefly here the derivation of the spin-flopping conditions. Exchange inversion and spin flopping are quite different phenomena and should not be confused. Besides the obvious fact that exchange inversion can take place in the absence of an externally applied field, the most important feature is that the F/AF transition can be induced by an external field applied along any crystallographic direction. Spin flopping, however, can occur only when the anisotropy of the antiferromagnetic state is positive and when the external field is applied parallel to the anisotropy field. The spin configuration in the "spin flopped" state is canted, and this state is stable only in the presence of the applied magnetic field. The state is not stabilized when the external field is applied perpendicular to the anisotropy field or when the anisotropy is negative.

The condition for an exchange-inversion transition has been treated thus far in the absence of external fields. However, an F/AF transition can be induced at temperatures below T_s by application of a sufficiently strong magnetic field. The condition for an F/AF transition in the presence of an applied field is given by equating the free energies for F and AF ordering, including the magnetic energy of the applied field,

$$F(\pi)=F(0)+2\mathbf{H}\cdot\mathbf{M} \quad (18)$$

from which is obtained the critical field for the F/AF transition

$$H_c^{\mathrm{AF/F}}=\frac{m}{M_0}\left\{\left(\frac{\rho m^2}{R}\right)^2\rho'+(c_T-c_c)\times\left[\rho+\frac{\rho'}{2}(c_T-c_c)\right]\right\}\left[1-\left(\frac{\rho' m^2}{R}\right)^2\right]^{-1}. \quad (19)$$

It can be verified that the derivative of Eq. (19) with respect to temperature and application of the magnetic form of the Clausius–Clapeyron equation gives Eq. (8). Also, it can be seen by direct substitution that the critical field vanishes when Eq. (5) is satisfied.

The anisotropy has been ignored in Eq. (18). Anisotropic exchange introduces anisotropy in the critical field for the F/AF transition. If this were the only contribution to the anisotropy, the calculated value of Δc_c gives a critical field perpendicular to the c axis about 2000 G larger than the field applied parallel for a given temperature. When all sources of anisotropy are included, one would expect the anisotropy of the critical field to be proportional to the magnetocrystalline anisotropy field. Such anisotropy in the critical field is mentioned by Bierstedt[10] in connection with his resis-

[9] R. B. Flippen, J. Appl. Phys. 34, 2026 (1963).
[10] P. E. Bierstedt, Phys. Rev. 132, 669 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tivity measurements. More detailed unpublished data of Bierstedt show that the lower critical field is parallel to the easy direction of the anisotropy. Thus, the lower critical field changes from parallel to perpendicular to the c axis for compositions with $T_s<200°$K, the temperature below which the anisotropy changes sign before the exchange inversion occurs.

The second-order quantity ρ'/R is small, and its inclusion contributes only additional complications. We shall neglect it in the succeeding calculation of spin flopping. Eq. (19) becomes

$$H_c^{\mathrm{AF/F}}=\frac{m}{M_0}\rho(c_T-c_c). \tag{20}$$

To obtain the spin-flopping conditions we again equate the free energies

$$F(\pi)=F(\theta)+2\mathbf{H}\cdot\mathbf{M}+2K\sin^2\theta. \tag{21}$$

Here the solution of Eq. (21) is anticipated by assuming that H is directed along the anisotropy field of the antiferromagnetic state and that the "spin-flopped" state is canted with the moments of the two sublattices making the same angle θ with respect to the applied field. Equation (21) may be rearranged to give

$$\cos\theta\left[\cos^3\theta+\left(\frac{m^2\rho(c_T-c_c)+K}{\rho^2m^4}R-1\right)\times\cos\theta+R\frac{H_c^{\mathrm{AF/C}}M}{\rho^2m^4}\right]-\frac{KR}{\rho^2m^4}=0, \tag{22}$$

where K is the anisotropy energy of the canted state. In order to obtain the critical field $H_c^{\mathrm{AF/C}}$ for spin flopping, the equilibrium value of θ must be determined by setting the torque equal to zero, $\partial F/\partial\theta=0$, where F is understood to contain the anisotropy and magnetic energies. Thus, the equilibrium condition

$$\cos^3\theta-\frac{1}{2}\left[1-\frac{K+2\rho(c_T-c_c)m^2}{2\rho^2m}R\right]\times\cos\theta+\frac{H_c^{\mathrm{AF/C}}M}{\rho^2m^4}-\frac{KR}{\rho^2m^4}R=0 \tag{23}$$

must be satisfied simultaneously with Eq. (22).

General solution of Eqs. (22) and (23) are complex, however, the essential features may be obtained by an approximate solution applicable to the experimental conditions of Flippen. Substitution of the field dependence of the F/AF transition and the observed critical field for spin flopping in Eq. (23) indicate that $\theta\cong90°$, and $\cos^3\theta$ can be neglected. Therefore, by solving Eq. (23) for $\cos\theta$ and substituting into Eq. (22), the spin-flopping condition is obtained

$$H_c^{\mathrm{AF/C}}=\frac{1}{M}\left(2\left\{K\left[\frac{\rho^2m^4}{R}-\rho(c_T-c_c)m^2\right]\right\}^{1/2}-K\left\{K\Big/\left[\frac{\rho^2m^4}{R}-\rho(c_T-c_c)m^2\right]\right\}^{1/2}\right). \tag{24}$$

When K is small compared to the magnetic strain ρ^2m^4/R, the second term of Eq. (24) may be neglected for any value of (c_T-c_c). Substitution of Eq. (20) gives the relation between the spin-flopping field and the field for exchange inversion

$$H_c^{\mathrm{AF/C}}=\frac{2}{M}\left\{K\left[\left(\frac{\rho^2m^4}{R}\right)+H_c^{\mathrm{F/AF}}M\right]\right\}^{1/2}. \tag{25}$$

The anisotropy field of the canted state is found to be of the order of 1000 Oe or 2×10^5 erg/cm^3 from analysis of the data of Flippen.

INTERMEDIATE STATE

Darnell *et al.*[6] and Austin *et al.*[5] show that at low transition temperatures a third magnetic state appears as an interruption in the F/AF transition. The foregoing model does not account for such behavior principally for two reasons: First, effectively only nearest-neighbor interactions are contained in the model, and there is no competition between exchange interactions for spin direction that might give rise to a helical structure. Second, the model assumes a Bravais lattice and even if the details of the exchange interactions were included, only proper helical structures would result. The spin configuration of the intermediate state must be represented by a mixture of two configurations,

$$\mathbf{S}_n=\mathbf{D}+\mathbf{U}\cos(\mathbf{k}\cdot\mathbf{r}_n)+\mathbf{V}\sin(\mathbf{k}\cdot\mathbf{r}_n), \tag{26}$$

a $k=0$ ferromagnetic component and a helical component commensurate with the lattice periodicity given by $k_z=2\pi/3c$. For $x<0.022$ in the composition $Mn_{2-x}Cr_xSb$ the intermediate state is stable at absolute zero. In order to account for such a spin configuration, one must treat the actual crystal structure of Mn_2Sb by the generalized Luttinger-Tisza method proposed by Lyons and Kaplan[11] and Lyons *et al.*[12] Preliminary calculations indicate that the spin eigenfunction given in Eq. (26) can lower the energy eigenvalues below the ferromagnetic and antiferromagnetic states. Details of this calculation will be published in a forthcoming paper.

DISCUSSION

The foregoing theory based on a highly simplified model of Mn_2Sb demonstrates considerable success in accounting for the behavior of the F/AF transition in

[11] D. H. Lyons and T. A. Kaplan, Phys. Rev. 120, 1580 (1960).
[12] D. H. Lyons, T. A. Kaplan, K. Dwight, and N. Menyuk, Phys. Rev. 126, 540 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$Mn_{2-x}Cr_xSb$. As pointed out by Bierstedt,[10] this transition is accompanied by hysteresis. For transitions near room temperature, the hysteresis is less than 1 deg of temperature. However, below about 200°K, hysteresis begins increasing sharply and reaches a value of about 20° for a transition occurring near the temperature of liquid nitrogen. First-order transitions to the ferromagnetic state in MnAs and the associated hysteresis have been investigated by Bean and Rodbell.[13] They point out that the maximum hysteresis arises when the spin system remains at the initial minimum of the free energy, even though it is not the lowest minimum, until the free energy is "down-hill-all-the-way." The measured hysteresis in MnAs is about 2/3 of this maximum value.

[13] C. P. Bean and D. S. Rodbell, Phys. Rev. 126, 104 (1962).

This situation does not hold for the F/AF transition in $Mn_{2-x}Cr_xSb$. Kittel has shown that the maximum hysteresis in a_T is $\Delta a_T = 2\rho m^2/R$. The free energy used in this treatment does not alter appreciably this expression. At room temperature, therefore, the expected hysteresis in Δa_T using ρ/R obtained previously is of the order of 10^{-2} Å, which may be transformed to a thermal hysteresis of about 50° by use of Fig. 2. This large discrepancy with the observed hysteresis indicates some very efficient mechanism nucleates the new state so that for transitions near room temperature the most stable state is nearly the equilibrium state. At low temperatures this nucleation mechanism must become less efficient, since the thermal hysteresis increases rapidly. A satisfactory model has not been obtained.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Magnetic Susceptibilities of Transition Elements in Host Crystals. II. Ni^{2+} in ZnO and CdS†

WILLIAM H. BRUMAGE AND CHUN C. LIN*
Department of Physics, University of Oklahoma, Norman, Oklahoma
(Received 10 December 1963)

The magnetic susceptibilities of Ni^{2+}-doped ZnO and CdS crystals have been measured along and perpendicular to the trigonal crystalline axes over the temperature range of 28–500°K. The magnetic susceptibilities of both crystals approach constant values at low temperature and decrease more rapidly with increasing temperature at $T>60$°K. An unusually large magnetic anisotropy is observed for Ni^{2+}:ZnO. The non-Curie behavior can be explained on the basis that the ground state of Ni^{2+} is nonmagnetic (A_1) and the temperature-dependent susceptibility arises mainly from the ions in the first two excited states (A_2 and E). By fitting the theoretical susceptibilities to the experimental values, the spin-orbit coupling constant of Ni^{2+}:ZnO is obtained as -175 ± 25 cm^{-1} and the trigonal field splitting of the $T_1[^3T(F)]$ state (the lowest T_1 state) as 100 ± 10 cm^{-1}. The corresponding quantities for Ni^{2+}:CdS are -170 ± 10 cm^{-1} and 10 ± 4 cm^{-1}. In both crystals the A_2 component of the $T_1[^3T(F)]$ state lies below the E level. The large reduction of the spin-orbit coupling constant from the free-ion value indicates a rather strong covalency between the Ni^{2+} ion and the ligands. Combination of the trigonal splittings of the $T_1[^3T_1(F)]$ state with those of $T_2[^3T_1(P)]$ observed in the optical spectra leads to a determination of the trigonal field parameters. The experimental values of the trigonal parameters are consistent with those calculated by using the point-charge model and assuming a local contraction of the lattice with a slightly larger contraction for the three anions off the trigonal axis than for the one on the axis.

I. INTRODUCTION

In a previous paper we have pointed out that the magnetic susceptibilities of V^{3+}-doped corundum may be used quite effectively for determining certain structural parameters of the crystals.[1] For example, the trigonal field splitting of the lowest $^3T_1$ level of V^{3+}:Al_2O_3 was determined from the Van Vleck temperature-independent susceptibility which was evaluated from the magnetic data in the temperature interval of 77–295°K, and the zero-field splitting of the ground state was obtained from the susceptibilities at low temperatures. By combining these results with those of electron spin resonance experiments and optical spectra,[2] the two trigonal field parameters and the spin-orbit coupling constant were estimated. The trigonal field parameters were found to be in reasonable agreement with the ones calculated from the empirical point-charge model. In this paper we shall report some magnetic measurements of Ni^{2+}:CdS and Ni^{2+}:ZnO. The susceptibility data make it possible to locate the first two excited states which hitherto have not been determined accurately.[3]

† Supported by the U. S. Office of Naval Research.
* Alfred P. Sloan Foundation Fellow.
[1] W. H. Brumage, C. R. Quade, and C. C. Lin, Phys. Rev. 131, 949 (1963).

[2] D. S. McClure, J. Chem. Phys. 36, 2757 (1962).
[3] An estimation of the energy of the first excited state of Ni^{2+}:ZnO has been made by intensity measurements. See R. Pappalardo, D. L. Wood, and R. C. Linares, Jr., J. Chem. Phys. 35, 1471 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Values of the spin-orbit coupling constant and the trigonal field parameters can then be obtained, and the latter are compared with the calculations of the point-charge model.

In both CdS and ZnO host crystals, a Ni^{2+} ion is situated in a tetrahedral site.[4,5] Since Ni^{2+} has an electron configuration of $(3d)^8$ and since the pattern of the energy level splitting due to a tetrahedral field is just the inverted form of that of an octahedral field, the ordering of the crystalline Stark levels of Ni^{2+}:CdS and V^{3+}:Al_2O_3 are identical if one ignores the effect of the trigonal distortion. However, the spin-orbit coupling constant of Ni^{2+} is considerably larger than that of V^{3+}. Thus, in Ni^{2+}:CdS the spin-orbit splittings are much greater than the trigonal terms, while these two effects are of comparable magnitude for Ni^{2+}:ZnO. It is convenient to consider the spin-orbit interaction as the major perturbation to Ni^{2+} in a cubic field and then introduce the trigonal terms as the second perturbation. Figure 1 shows the energy level diagram of a Ni^{2+} ion in tetrahedral site.

The ground state A_1 is nonmagnetic, i.e., the average magnetic moment for this state is zero. At very low temperature the susceptibility should approach a constant value which is equal to the sum of the diamagnetism of the host crystal and the Van Vleck temperature-independent susceptibility. If we choose the trigonal axis as the z axis, it is easily shown that the matrix elements of μ_x and μ_y connect an A_1 state with only the E states, while μ_z has elements between A_1 and A_2 only. It follows that for an external magnetic field along the z axis, the Van Vleck term is dictated primarily by the first A_2 level (in the trigonal field) while in the case of a perpendicular field, the lowest E state gives the major contribution to the temperature-independent paramagnetic susceptibility. As the temperature is increased, the excited states begin to receive appreciable population and the susceptibilities then depart from the low-temperature asymptotic values. Because of this strong dependence of χ_z and χ_x on temperature, the positions of the first A_2 and the first E states can be determined quite accurately. In the case of Ni^{2+}:CdS, we are able to obtain from the magnetic data a trigonal splitting of the lowest A_2–E pair as small as 10 cm^{-1}. Furthermore, since both the first excited A_2 and E levels are a few hundred cm^{-1} above the ground state, direct determination of the location of these levels is difficult by other means and the magnetic measurements are particularly effective for this purpose.

II. EXPERIMENT

The crystals of Ni^{2+}:CdS were furnished by Dr. T. L. Estle of Texas Instrument Inc., and Ni^{2+}:ZnO by Dr. H. A. Weakleim of the RCA Laboratories. The measurements of the magnetic susceptibility were made by means of the Faraday balance used in our previous work.[1] In Tables I and II are shown the magnetic susceptibilities of Ni^{2+}:CdS and Ni^{2+}:ZnO crystals along and perpendicular to the trigonal axis over the temperature range of 28–500°K. The concentrations of the nickel ions in ZnO and CdS are 0.076 and 0.84%, respectively, as determined from the analysis of the magnetic data. The accuracy of the measured susceptibility in ZnO (±4%) is somewhat lower than in CdS (±2%) because of the smaller concentration of Ni^{2+} in the former.

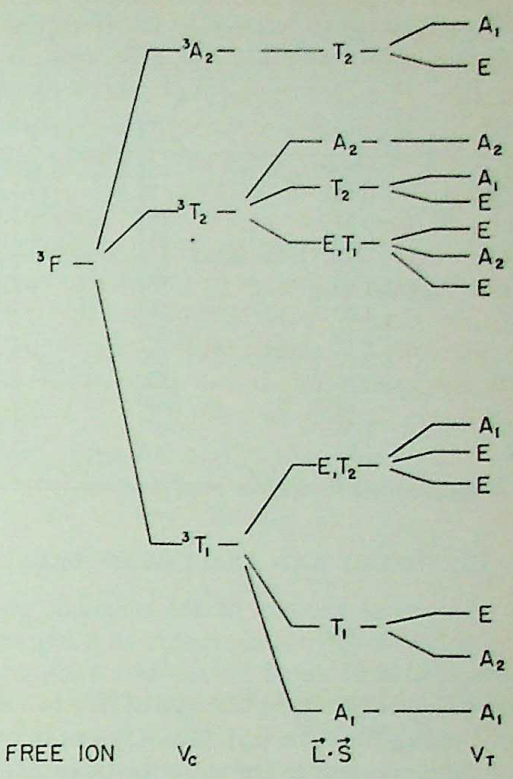


FIG. 1. Energy levels (^3F) of Ni^{2+} in a trigonal field.

TABLE I. Magnetic susceptibilities (per gram sample) of Ni^{2+}-doped CdS crystal. Concentration: 0.84%.

T (°K)	$\chi_{\parallel}\times10^7$ (cgs-emu)	T (°K)	$\chi_{\perp}\times10^7$ (cgs-emu)
77.3	8.30	77.3	7.90
94	7.75	86	7.60
108	7.12	101	7.18
128	6.32	121	6.35
154	5.40	134	5.85
174	4.68	164	4.80
202	3.86	184	4.16
217	3.51	203	3.72
232	3.11	219	3.48
255	2.63	249	2.77
296	1.98	294	1.98
391	0.90	389	0.80
489	0.21	448	0.37

[4] R. W. G. Wyckoff, *Crystal Structure* (Interscience Publishers, Inc., New York, 1948), Vol. 1; G. Heiland, E. Mollwo, and F. Stockmann, Solid State Phys. 8, 195 (1959).
[5] H. A. Weakliem, J. Chem. Phys. 36, 2117 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Magnetic susceptibilities (per gram sample) of Ni^{2+}-doped ZnO crystal. Concentration: 0.076%.

T (°K)	$\chi_{\parallel}\times10^7$ (cgs-emu)	T (°K)	$\chi_{\perp}\times10^7$ (cgs-emu)
28.0	0.00	28.0	−1.85
77.3	−0.27	77.3	−1.85
97	−0.53	107	−1.85
108	−0.75	114	−1.95
130	−1.08	118	−2.00
140	−1.30	151	−2.10
203	−1.90	167	−2.13
253	−2.16	201	−2.35
273	−2.22	250	−2.46
300	−2.33	300	−2.57
335	−2.42	330	−2.73
403	−2.66	401	−2.86
501	−2.90	501	−2.90

III. THEORY AND ANALYSIS OF DATA

The theoretical analysis of the magnetic susceptibility of a (d^2) or (d^8) atomic system in a trigonal field has been treated in detail for the $V^{3+}:Al_2O_3$ crystals.[1] As pointed out in Sec. I, for the case of Ni^{2+} in a slightly distorted tetrahedral site, it is advantageous to consider the spin-orbit coupling as the major perturbation on the perfect tetrahedral field and the trigonal distortion as a minor one. Thus, a slight modification will be made to the theory presented in Ref. 1.

The Hamiltonian for a Ni^{2+} ion is written as

$$H=H_0+V_c+\lambda\mathbf{L}\cdot\mathbf{S}+V_t+\mu_0(\mathbf{L}+g_s\mathbf{S})\cdot\mathcal{H}. \quad (1)$$

The various members in the right-hand side of Eq. (1) represent the Hamiltonian of a free Ni^{2+} ion, the cubic (tetrahedral) crystalline field, the spin-orbit interaction term, the trigonal field, and the interaction with the external field $\mathcal{H}$. As was done before, the two crystal potential terms are taken as[6]

$$V_c=\sum_{i=1,2} r_i^4 A_4^0\{Y_{4,0}(\theta_i\phi_i)+(10/7)^{1/2}\times[Y_{4,3}(\theta_i\phi_i)-Y_{4,-3}(\theta_i\phi_i)]\}, \quad (2)$$

$$V_t=\sum_{i=1,2}\{r_i^2 B_2^0 Y_{2,0}(\theta_i\phi_i)+r_i^4 B_4^0 Y_{4,0}(\theta_i\phi_i)\}. \quad (3)$$

Here the summation is to be extended over the two $(3d)$ holes rather than electrons. The crystal-field parameters are defined as

$$\beta=-A_4^0\langle r^4\rangle/14\sqrt{\pi},$$
$$\gamma=-B_4^0\langle r^4\rangle/42\sqrt{\pi},$$
$$\tau=-(\sqrt{5})B_2^0\langle r^2\rangle/140\sqrt{\pi},$$

and, in the case of a pure cubic field β is equal to $\frac{2}{3}Dq$.

Let us first consider the part of the Hamiltonian

$$H'=H_0+V_c+\lambda\mathbf{L}\cdot\mathbf{S}. \quad (4)$$

[6] See, for example, D. S. McClure, Solid State Phys. 9, 399 (1959).

The lowest configuration $(3d)^8$ of a Ni^{2+} ion gives rise to the terms ^3F, ^1D, ^3P, ^1G, and ^1S which are split by the cubic field, e.g.,

$$^3F\rightarrow{}^3T_1+{}^3T_2+{}^3A_2, \quad ^3P\rightarrow{}^3T_1. \quad (5)$$

The spin-orbit coupling causes further splitting of the levels as

$$^3T_1\rightarrow A_1+E+T_1+T_2,$$
$$^3T_2\rightarrow A_2+E+T_1+T_2, \quad (6)$$
$$^3A_2\rightarrow T_2.$$

The eigenfunctions of H_0 corresponding to the various terms of the $(3d)^8$ configuration are taken as the usual Clebsch-Gordan type combinations[7] of one-electron orbitals, and the multiplet separations are treated as parameters to be determined from the optical spectra. From the eigenfunctions of H_0 we can construct the cubic-field functions according to Table IV of Ref. 1. These functions are characterized by the irreducible representations of the cubic group to which they belong. Since V_c has matrix elements connecting $^3T_1(^3F)$ and $^3T_1(^3P)$, a transformation

$$\begin{matrix}{}^3F\\{}^3P\end{matrix}\begin{pmatrix} a_1 & a_2\\ -a_2 & a_1\end{pmatrix} \quad (7)$$

must be applied to these two states in order to diago-

TABLE III. Spin-orbit functions of ^3F and ^3P.

^3F

$^3T_1$:
$\phi(A_1,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)-\Psi({}^3T_1,0,0)+\Psi({}^3T_1,-,+1)\}/\sqrt{3}$
$\phi(T_1,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)-\Psi({}^3T_1,-,+1)\}/\sqrt{2}$
$\phi(T_1,{}^3T_1,+)=\{\Psi({}^3T_1,+,0)-\Psi({}^3T_1,0,+1)\}/\sqrt{2}$
$\phi(T_1,{}^3T_1,-)=\{-\Psi({}^3T_1,-,0)+\Psi({}^3T_1,0,-1)\}/\sqrt{2}$
$\phi(E,{}^3T_1,+)=\{-\Psi({}^3T_1,-,-1)+\Psi({}^3T_1,+,0)+\Psi({}^3T_1,0,+1)\}/\sqrt{3}$
$\phi(E,{}^3T_1,-)=\{+\Psi({}^3T_1,-,-1)+\Psi({}^3T_1,+,0)+\Psi({}^3T_1,0,+1)\}/\sqrt{3}$
$\phi(T_2,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)+2\Psi({}^3T_1,0,0)+\Psi({}^3T_1,-,+1)\}/\sqrt{6}$
$\phi(T_2,{}^3T_1,+)=\{2\Psi({}^3T_1,+,+1)-\Psi({}^3T_1,-,0)-\Psi({}^3T_1,0,-1)\}/\sqrt{6}$
$\phi(T_2,{}^3T_1,-)=\{-2\Psi({}^3T_1,-,-1)-\Psi({}^3T_1,+,0)-\Psi({}^3T_1,0,+1)\}/\sqrt{6}$

$^3T_2$:
$\phi(A_2,{}^3T_2,0)=\{-\Psi({}^3T_2,0,0)-\Psi({}^3T_2,-,+1)+\Psi({}^3T_2,+,-1)\}/\sqrt{3}$
$\phi(T_2,{}^3T_2,0)=\{\Psi({}^3T_2,+,-1)+\Psi({}^3T_2,-,+1)\}/\sqrt{2}$
$\phi(T_2,{}^3T_2,+)=\{\Psi({}^3T_2,+,0)-\Psi({}^3T_2,0,+1)\}/\sqrt{2}$
$\phi(T_2,{}^3T_2,-)=\{\Psi({}^3T_2,0,-1)+\Psi({}^3T_2,-,0)\}/\sqrt{2}$
$\phi(E,{}^3T_2,+)=\{\Psi({}^3T_2,-,-1)+\Psi({}^3T_2,0,+1)+\Psi({}^3T_2,+,0)\}/\sqrt{3}$
$\phi(E,{}^3T_2,-)=\{\Psi({}^3T_2,+,+1)+\Psi({}^3T_2,0,-1)-\Psi({}^3T_2,-,0)\}/\sqrt{3}$
$\phi(T_1,{}^3T_2,0)=\{2\Psi({}^3T_2,0,0)-\Psi({}^3T_2,-,+1)+\Psi({}^3T_2,+,-1)\}/\sqrt{6}$
$\phi(T_1,{}^3T_2,+)=\{2\Psi({}^3T_2,+,+1)-\Psi({}^3T_2,0,-1)-\Psi({}^3T_2,-,0)\}/\sqrt{6}$
$\phi(T_1,{}^3T_2,-)=\{-2\Psi({}^3T_2,-,-1)-\Psi({}^3T_2,0,+1)-\Psi({}^3T_2,+,0)\}/\sqrt{6}$

$^3A_2$:
$\phi(T_2,{}^3A_2,0)=\Psi({}^3A_2,0)$
$\phi(T_2,{}^3A_2,+)=\Psi({}^3A_2,+1)$
$\phi(T_2,{}^3A_2,-)=\Psi({}^3A_2,-1)$

^3P

$^3T_1$:
$\phi(A_1,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)-\Psi({}^3T_1,0,0)+\Psi({}^3T_1,-,+1)\}/\sqrt{3}$
$\phi(T_1,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)-\Psi({}^3T_1,-,+1)\}/\sqrt{2}$
$\phi(T_1,{}^3T_1,+)=\{\Psi({}^3T_1,+,0)-\Psi({}^3T_1,0,+1)\}/\sqrt{2}$
$\phi(T_1,{}^3T_1,-)=\{\Psi({}^3T_1,0,-1)-\Psi({}^3T_1,-,0)\}/\sqrt{2}$
$\phi(E,{}^3T_1,+)=\{-\Psi({}^3T_1,-,-1)+\Psi({}^3T_1,+,0)+\Psi({}^3T_1,0,+1)\}/\sqrt{3}$
$\phi(E,{}^3T_1,-)=\{\Psi({}^3T_1,+,+1)+\Psi({}^3T_1,0,-1)+\Psi({}^3T_1,-,0)\}/\sqrt{3}$
$\phi(T_2,{}^3T_1,0)=\{\Psi({}^3T_1,+,-1)+2\Psi({}^3T_1,0,0)+\Psi({}^3T_1,-,+1)\}/\sqrt{6}$
$\phi(T_2,{}^3T_1,-)=\{-2\Psi({}^3T_1,-,-1)-\Psi({}^3T_1,+,0)-\Psi({}^3T_1,0,1)\}/\sqrt{6}$
$\phi(T_2,{}^3T_1,+)=\{2\Psi({}^3T_1,+,+1)-\Psi({}^3T_1,0,-1)-\Psi({}^3T_1,-,0)\}/\sqrt{6}$

[7] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, London, 1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

nalize H_0+V_c. This results in a mixing between the F and P states. When the spin-orbit coupling is taken into consideration, the nine functions associated with the $^3T_1$ (also $^3T_2$) must be regrouped in such a way as to exhibit the decomposition shown in (6). The symbol $\phi(T_2,{}^3T_1,+)$ is used to denote a new function obtained in this manner corresponding to the T_2 representation and originating from the $^3T_1$ state. The last symbol inside the parenthesis specifies a particular component of the degenerate state. The functions ϕ are listed in Table III. In the representation with ϕ as the basis, H' is not completely diagonal since $\mathbf{L}\cdot\mathbf{S}$ does connect members of the $^3T_1$ manifold with those of $^3T_2$.

When the trigonal field is introduced, the T_1 and T_2 split into A_2+E and A_1+E, respectively. The energy matrix (in the absence of external field) then factorizes into blocks corresponding to A_1, A_2, and E. For magnetic susceptibilities below 500°K, the highly excited states are insignificant, so we will consider only the manifold associated with the ^3F state (with F-P mixing included). The energy matrix is shown in Table IV. Any mixing of these states with higher excited states can arise only from the spin-orbit coupling or from V_t and are too small to affect the susceptibilities.

A. Ni^{2+}:CdS

The optical studies reported by Weakliem[5] suggest that the environment of the Ni^{2+} in CdS is very nearly cubic because of the apparent isotropy of the spectrum. The same conclusion can also be reached from the small observed magnetic anisotropy (Table I). It is therefore convenient to diagonalize first the Hamiltonian H', defined by Eq. (3), and subsequently treat V_t by first-order perturbation. To diagonalize H' we shall take the values of β and $\triangle_P$ (the spacing between ^3F and ^3P for the limiting free ion) as 257 and 8550 cm^{-1}, respectively, according to the results of Weakliem.[5] This gives the F-P mixing coefficients as

$$a_1=0.991\,,\quad a_2=0.138.$$

The wave functions which diagonalize H' are now labeled by the irreducible representations of the cubic group. Thus, a wave function of a particular symmetry species can be expressed as linear combinations of the functions listed in Table III, e.g.,

$$\psi(T_2,+)=a\phi(T_2,{}^3T_1,+)+b\phi(T_2,{}^3T_2,+). \qquad (8)$$

When the trigonal potential is introduced, these functions may still be used as the approximate wave function of $H'+V_t$, and the effect of V_t is to produce splittings or shifts of the energy levels. The wave functions and energies of the lowest nine states which are associated with the $^3T_1(F)$ group are determined in this manner.

With the wave functions for these nine states, the magnetic susceptibilities have been calculated by means of the standard procedure[8] and are expressed in terms of the energy levels of the low-lying states and the matrix elements of $\mathbf{L}$ and $\mathbf{S}$ between these states. The formulas for $\chi_{\|}$ and $\chi_{\perp}$ are

$$\chi_{\|}=2NB\Big[\frac{|(0|\mu_z|1)|^2}{\Delta_1}+\exp(-\Delta_1/kT)\times\Big\{-\frac{|(0|\mu_z|1)|^2}{\Delta_1}+\frac{|(1|\mu_z|4)|^2}{\Delta_4-\Delta_1}\Big\}+2\exp(-\Delta_2/kT)\times\Big\{\frac{|(2|\mu_z|2)|^2}{2kT}+\sum_{i\neq 2}\frac{|(2|\mu_z|i)|^2}{\Delta_i-\Delta_2}\Big\}+2\exp(-\Delta_3/kT)\times\Big\{\frac{|(3|\mu_z|3)|^2}{2kT}+\sum_{i\neq 3}\frac{|(3|\mu_z|i)|^2}{\Delta_i-\Delta_3}\Big\}+\exp(-\Delta_4/kT)\times\Big\{\frac{|(1|\mu_z|4)|^2}{\Delta_1-\Delta_4}\Big\}+2\exp(-\Delta_5/kT)\times\Big\{\frac{|(5|\mu_z|5)|^2}{2kT}+\sum_{i\neq 5}\frac{|(5|\mu_z|i)|^2}{\Delta_i-\Delta_5}\Big\}\Big], \qquad (9)$$

$$\chi_{\perp}=4NB\sum_i\Big[\frac{|(0|\mu_x|i)|^2}{\Delta_i}+\exp(-\Delta_1/kT)\times\Big\{\sum_i\frac{|(1|\mu_x|i)|^2}{\Delta_i-\Delta_1}\Big\}+\exp(-\Delta_2/kT)\times\Big\{\frac{|(2|\mu_x|2)|^2}{2kT}+\sum_{j\neq 2}\frac{|(2|\mu_x|j)|^2}{\Delta_j-\Delta_2}\Big\}+\exp(-\Delta_3/kT)\times\Big\{\frac{|(3|\mu_x|3)|^2}{2kT}+\sum_{j\neq 3}\frac{|(3|\mu_x|j)|^2}{\Delta_j-\Delta_2}\Big\}+\exp(-\Delta_4/kT)\times\Big\{\sum_i\frac{|(4|\mu_x|i)|^2}{\Delta_i-\Delta_4}\Big\}+\exp(-\Delta_5/kT)\times\Big\{\frac{|(5|\mu_x|5)|^2}{2kT}+\sum_{j\neq 5}\frac{|(5|\mu_x|j)|^2}{\Delta_j-\Delta_5}\Big\}\Big],$$

where

$$B^{-1}=1+\exp(-\Delta_1/kT)+2\exp(-\Delta_2/kT)+2\exp(-\Delta_3/kT)+\exp(-\Delta_4/kT)+2\exp(-\Delta_5/kT),$$

and

$$i=2,\,3,\,5;$$
$$j=0,\,1,\,2,\,3,\,4,\,5.$$

In the above equations the energy spacing between the kth level and the ground level is represented by Δ_k and the numbers 0, 1, 2, 3, 4, 5 represent the A_1, T_1A_2,

[8] J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities* (Oxford University Press, London, 1932), p. 182.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IV. Energy matrix.

A_1:	$^3T_1A(^3F)$	$^3T_1T_2(^3F)$	$^3T_2T_2(^3F)$	$^3A_2T_2(^3F)$	$^3T_1A(^3P)$	$^3T_1T_2(^3P)$
	$-9\beta+3\lambda-7\gamma$	$\sqrt{2}(-10\gamma-\tau)$	$-\frac{1}{3}(30)^{1/2}(4\gamma-\tau)$	$-\frac{2}{3}(15)^{1/2}(\gamma-2\tau)$	$-6\beta-(14/3)\gamma$	$(5/3)\sqrt{2}\gamma-8\sqrt{2}\tau$
		$-9\beta-\frac{3}{2}\lambda+3\gamma+\tau$	$\frac{1}{2}(15)^{1/2}\lambda-\frac{1}{3}(15)^{1/2}(4\gamma-\tau)$	$-\frac{2}{3}(30)^{1/2}(\gamma-2\tau)$	$(5/3)\sqrt{2}\gamma-8\sqrt{2}\tau$	$-6\beta-(11/3)\gamma+(8/3)\tau$
			$3\beta-\frac{1}{2}\lambda-\gamma-5\tau$	$2\sqrt{2}\lambda$	$2(\frac{5}{6})^{1/2}(3\gamma-8\tau)$	$(5/3)^{1/2}(3\gamma-8\tau)$
				$18\beta+14\gamma$	$\sqrt{3}(4\gamma-20\tau)$	$(6)^{1/2}(-4\gamma+15\tau)$
					$\triangle_p-2\lambda$	$-14\sqrt{2}\tau$
						$\triangle_p+\lambda+14\tau$
A_2:	$^3T_1T_1(^3F)$	$^3T_2A(^3F)$	$^3T_2T_1(^3F)$			$^3T_1T_1(^3P)$
	$-9\beta+\frac{3}{2}\lambda-17\gamma-\tau$	$-\frac{1}{3}(30)^{1/2}(4\gamma-\tau)$	$\frac{1}{2}(15)^{1/2}\lambda-\frac{1}{3}(15)^{1/2}(4\gamma-\tau)$			$-6\beta-3\gamma-8\tau$
		$3\beta-\lambda+(7/3)\gamma$	$\frac{1}{3}\sqrt{2}(-10\gamma-15\tau)$			$-\frac{1}{3}(30)^{1/2}(-3\gamma-8\tau)$
			$3\beta+\frac{1}{2}\lambda+(7/3)\gamma$			$-\frac{1}{3}(15)^{1/2}(-3\gamma-8\tau)$
						$\triangle_p-\lambda-14\tau$
E:	$^3T_1T_1(^3F)$	$^3T_1E(^3F)$	$^3T_1T_2(^3F)$	$^3T_2T_2(^3F)$		$^3T_2E(^3F)$
	$-9\beta+\frac{3}{2}\lambda-2\gamma+\frac{1}{2}\tau$	$(6)^{1/2}(-5\gamma-\frac{1}{2}\tau)$	$+\sqrt{3}(5\gamma-\frac{1}{2}\tau)$	$-2(5)^{1/2}\gamma+\frac{1}{2}(5)^{1/2}\tau$		$-\frac{1}{6}(30)^{1/2}(4\gamma-\tau)$
		$-9\beta-\frac{3}{2}\lambda-7\gamma$	$\sqrt{2}(-5\gamma-\frac{1}{2}\tau)$	$\frac{1}{6}(30)^{1/2}(-4\gamma+\tau)$		$\frac{3}{2}(5)^{1/2}\lambda$
			$-9\beta-\frac{3}{2}\lambda-12\gamma-\frac{1}{2}\tau$	$\frac{1}{2}(15)^{1/2}\lambda+\frac{1}{6}(15)^{1/2}(4\gamma-\tau)$		$\frac{1}{2}(10)^{1/2}(4\gamma-\tau)$
				$3\beta-\frac{1}{2}\lambda+4\gamma+\frac{5}{2}\tau$		$\frac{1}{6}(6)^{1/2}(-10\gamma-15\tau)$
						$3\beta+\frac{1}{2}\lambda+(7/3)\gamma$
E:	$^3T_2T_1(^3F)$	$^3A_2T_2(^3F)$	$^3T_1T_1(^3P)$	$^3T_1E(^3P)$		$^3T_1T_2(^3P)$
	$\frac{1}{2}(15)^{1/2}\lambda+\frac{(15)^{1/2}}{6}(4\gamma-\tau)$	$(10)^{1/2}(-\gamma+2\tau)$	$-6\beta-(11/2)\gamma+4\tau$	$(6)^{1/2}(-\frac{5}{6}\gamma-4\tau)$		$\sqrt{3}(\frac{5}{6}\gamma+4\tau)$
	$\frac{1}{2}(10)^{1/2}(4\gamma-\tau)$	$\frac{(15)^{1/2}}{3}(2\gamma-4\tau)$	$(6)^{1/2}(-\frac{5}{6}\gamma-4\tau)$	$-6\beta-(14/3)\gamma$		$-\frac{5}{6}\sqrt{2}\gamma-4\sqrt{2}\tau$
	$(5)^{1/2}(2\gamma-\frac{1}{2}\tau)$	$\frac{(30)^{1/2}}{6}(-2\gamma+4\tau)$	$\sqrt{3}(\frac{5}{6}\gamma+4\tau)$	$\sqrt{2}(-\frac{5}{6}\gamma-4\tau)$		$-6\beta+\frac{1}{6}\sqrt{2}(-23\gamma-24\tau)$
	$\frac{1}{6}\sqrt{3}(10\gamma+15\tau)$	$2\sqrt{2}\lambda$	$\frac{1}{2}(5)^{1/2}(3\gamma+8\tau)$	$(\frac{5}{6})^{1/2}(3\gamma+8\tau)$		$(5/12)^{1/2}(-3\gamma-8\tau)$
	$\frac{1}{6}\sqrt{2}(-10\gamma-15\tau)$	0	$(\frac{5}{6})^{1/2}(3\gamma+8\tau)$	0		$\frac{1}{2}(10)^{1/2}(-3\gamma-8\tau)$
	$3\beta+\frac{1}{2}\lambda+\frac{2}{3}\gamma-\frac{5}{2}\tau$	0	$(5/12)^{1/2}(-3\gamma-8\tau)$	$\frac{1}{2}(10)^{1/2}(-3\gamma-8\tau)$		$\frac{1}{2}(5)^{1/2}(-3\gamma-8\tau)$
		$18\beta+14\gamma$	$\sqrt{2}(6\gamma-30\tau)$	$\sqrt{3}(-4\gamma+20\tau)$		$(6)^{1/2}(2\gamma-10\tau)$
			$\triangle_p-\lambda+7\tau$	$-7(6)^{1/2}\tau$		$+7\sqrt{3}\tau$
				$\triangle_p+\lambda$		$-7\sqrt{2}\tau$
						$\triangle_p+\lambda-7\tau$

T_1E, EE, T_2A_1, and T_2E levels, respectively. Here the matrix elements of **L** and **S** are to be evaluated using the wave functions of the type of Eq. (8). These elements can be written as functions of the mixing coefficients as defined in Eq. (8) which, in turn, depends on the various crystal field parameters.[9] The algebraic work, which is quite lengthy, will not be reproduced

[9] W. H. Brumage, Ph.D. thesis, University of Oklahoma, 1964 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

here and the readers are referred to Ref. 9 for the mathematical details.

The excited states beyond the $^3T_1(F)$ manifold are at least about 3000 cm^{-1} above the ground state, so they can be ignored as far as the magnetic susceptibility is concerned. It was found that the theoretical susceptibilities at temperatures below 500°K depend on λ and a parameter K which is defined as

$$K=-30\gamma-3\tau, \tag{10}$$

but is practically independent of the individual values of γ and τ. The calculated susceptibilities can be fitted to the experimental data by choosing $\lambda=-170$ cm^{-1}, and $K=-20$ cm^{-1} (see Fig. 2). These give the positions of the first two excited states (A_2 and E) as 230 and 240 cm^{-1}, respectively. It is rather remarkable to note that one is able to determine a splitting as small as 10 cm^{-1} from the magnetic measurements. This occurs because the magnetic anisotropy depends quite sensitively on this A_2-E splitting. When the uncertainty of the experimental data is taken into consideration, the trigonal splitting is determined as 10 ± 4 cm^{-1}. Likewise, the accuracy of λ and K is estimated to be

$$\lambda=-170\pm10 \text{ cm}^{-1},$$
$$K=-20\pm8 \text{ cm}^{-1}. \tag{11}$$

The determination of the individual values of γ and τ will be discussed in Sec. IV.

B. Ni^{2+}:ZnO

The rather large magnetic anisotropy of Ni^{2+}:ZnO (Fig. 3) indicates that the trigonal field here is considerably greater than that in Ni^{2+}:CdS. In fact, from the susceptibility data we find that the spin-orbit coupling constant and the trigonal field parameters are of comparable magnitude. The approximate method for determining the energy levels and wave functions used in the previous section is no longer applicable here; rather one must solve the entire 21×21 secular equation as given in Table III. The amount of F-P mixing in this case is, however, quite small and can be neglected. With a given set of values of λ, γ, and τ, one can solve the secular equation numerically and calculate $\chi_{\parallel}$ and $\chi_{\perp}$ according to Eq. (9). The wave functions for the lowest nine states are given in Ref. 9. Again the calculated susceptibilities are governed chiefly by λ and K, and by fitting the theoretical values to the experimental data, we found

$$\lambda=-175\pm25 \text{ cm}^{-1},$$
$$K=-225\pm25 \text{ cm}^{-1}. \tag{12}$$

Figure 3 shows the comparison between the theoretical and observed susceptibilities. The energies of the first two excited states (A_2 and E) are 160 and 260 cm^{-1}.

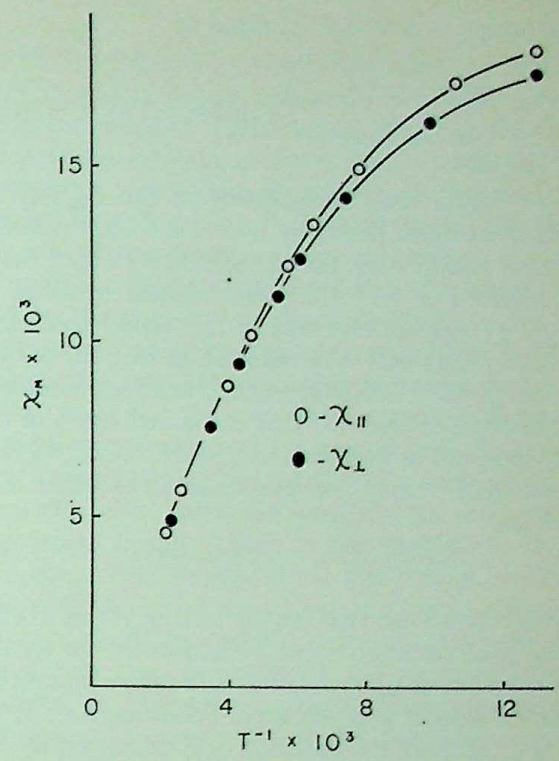


FIG. 2. Magnetic susceptibilities (per gm-ion) of Ni^{2+}:CdS from 77–500°K.

IV. DISCUSSION

From the optical spectrum, Weakliem[5] has determined the trigonal splitting of the level $T_2[^3T_1(P)]$ of Ni^{2+}:ZnO as 41 cm^{-1}. By combining this splitting with the parameter K given in Eq. (10), the two trigonal

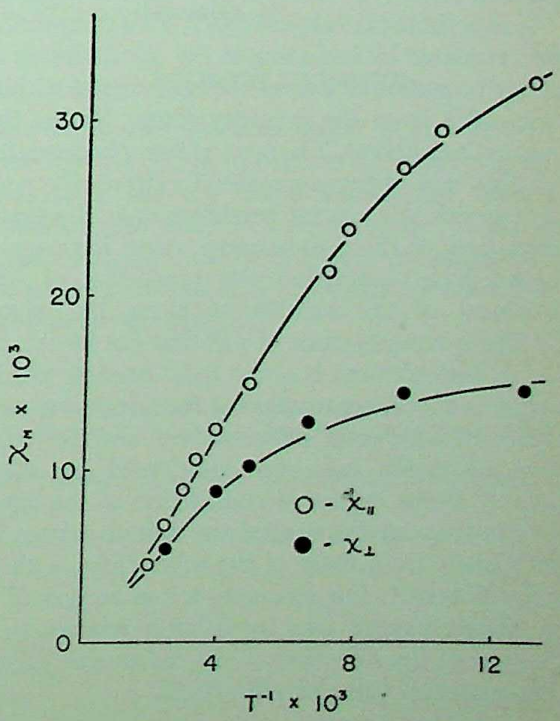


FIG. 3. Magnetic susceptibilities (per gm-ion) of Ni^{2+}:ZnO from 77–500°K.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

field parameters can be determined as

$$\left.\begin{aligned}\gamma &= 7.5\ \text{cm}^{-1}\\ \tau &= -1.4\ \text{cm}^{-1}\end{aligned}\right\},\quad \text{for } Ni^{2+}:ZnO. \tag{13}$$

Here we found that the splitting of the $T_2[^3T_1(P)]$ state is determined primarily by τ. On the other hand, K depends mainly on γ, the contribution of τ on K being quite small. For Ni^{2+}:CdS, the trigonal splitting of $T_2[^3T_1(P)]$ was not observed in the optical spectrum, presumably too small to be resolved. In fact, the optical spectrum of Ni^{2+}:CdS shows nearly complete isotropy. If, however, it is assumed that the contribution of the τ term in K can be neglected also for Ni^{2+}:CdS, as is in the case of Ni^{2+}:ZnO, one can make an estimate of γ from K as

$$\gamma \cong 0.67\ \text{cm}^{-1},\quad \text{for } Ni^{2+}:CdS. \tag{14}$$

It has been shown that for the case of $V^{3+}:Al_2O_3$ the trigonal field parameters can be accounted for by the empirical point-charge model.[1] It is therefore interesting to make a similar analysis here. Following Ref. 1, we shall first compare the values of γ/β calculated by the point-charge model with the experimental results. The use of γ/β, rather than γ, gives a better test of the point-charge model, since the former is independent of $\langle r^4\rangle$ as well as the effective charge assigned to the ions of the host crystals to account for the departure from purely ionic bonding. Let us first assume that the impurity ion is situated exactly at the site of Zn^{2+}. From the crystal structure data[4] and using only the four nearest-neighbor oxygen ions, we obtain from the point-charge calculation $\gamma/\beta=0.04$ for Ni^{2+}:ZnO as compared with the experimental value of 0.027. This computation has been repeated by including in the calculation of the crystalline potential all ions of the host crystal within a radius of 20 Å from the impurity center (rather than four nearest neighbors). The ratio γ/β now becomes 0.05 which does not differ substantially from the value calculated with the nearest neighbors only. To explain the deviation of the point-charge value from experiment, we have considered the possibility of small displacement of the substituent along the trigonal axis.[1,2] The minimum value of γ/β that can be obtained by such a displacement is about 0.04. Another possible distortion is a local contraction of the lattice due to the difference in ionic radii between Ni^{2+} and Zn^{2+}. This type of distortion has been considered by several authors.[5,10] If the degree of contraction of the oxygen ion on the trigonal axis (called the unique oxygen ion) differs slightly from those of the other three, a change of γ/β will result. For example, a contraction of the three off-axis oxygen ions by 0.028 Å relative to the unique ion in the direction of the cation site will give the experimental value of $\gamma/\beta=0.027$.

The calculation of τ from the point-charge model is considerably more difficult, as it is necessary to introduce an additional assumption of using the self-consistent field atomic wave function of Ni^{2+} to evaluate $\langle r^2\rangle$ which is not always valid.[11] Also, the convergence of this parameter in summing over the lattice is rather poor. We have carried out the summation to 30 A and the results still fluctuate considerably. However, we are able to obtain a rough estimate of τ as 1 cm^{-1} for the case of zero displacement of the Ni^{2+} from the cation site and no distortion of the lattice. With the distorted lattice as used in the previous paragraph (and no displacement of the impurity ions), the value of τ becomes -7 cm^{-1} which is of the same sign as the experimental value. In view of the assumptions involved in the calculation of τ, it is not reasonable to expect anything more than a qualitative agreement with experiment, and thus the point-charge calculation of this parameter will not be pursued further at this stage.

In the case of Ni^{2+}:CdS, the point-charge model gives, for zero displacement and no distortion, $\gamma/\beta=0.018$ which is much larger than the experimental value. We have used only the nearest-neighbor anions for this calculation since the inclusion of ions at lattice sites further away, as we have seen earlier, will not change this parameter greatly. From the apparent isotropy of the optical spectrum[5] as well as the smallness of γ/β, we note that the crystalline field around the Ni^{2+} in CdS has considerably smaller trigonal components than is inferred from the crystal structure of CdS. Like the case of Ni^{2+}:ZnO, the experimental value of γ/β cannot be explained by the displacement of the substituent alone. On the other hand, the picture of local distortion where the lattice relaxes toward the impurity ion in such a manner as to restore back to a nearly perfect tetrahedral configuration, does account for the small apparent anisotropy. Indeed, the experimental value of γ/β corresponds to a contraction of the three off-axis anions through 0.02 Å relative to the fourth one. The idea of local lattice contraction in Ni^{2+}:CdS was suggested by Weakliem in his optical studies.[5]

Thus, in both Ni^{2+}:ZnO and Ni^{2+}:CdS, the experimental values of the trigonal field parameters are consistent with those calculated from the empirical point-charge model[1] with a local distortion of the lattice such that the three anions of the trigonal axis relax toward the Ni^{2+} slightly more than does the one on the axis. Of course, the point-charge model is a highly idealized picture. When it is applied to crystals with an appreciable amount of covalency, a "displacement" of the anion may result from an actual movement of the lattice site as well as a shift of the electron cloud due to a change of the degree of covalency. In our application here no sharp distinction can be made between these two effects and the term "lattice distortion" really

[10] R. S. Title, Phys. Rev. 131, 2503 (1963).

[11] D. S. McClure, J. Chem. Phys. 38, 2289 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

refers to a movement of the effective center of charge of the anion.

The spin-orbit coupling constants of Ni^{2+} in both CdS and ZnO are about one-half of the free-ion value (-340 cm^{-1}). They are also substantially smaller than the spin-orbit constants of Ni^{2+} in octahedral field, e.g., -245 cm^{-1} for Ni^{2+}:MgO,[12] and -250 cm^{-1} (or -270) for Ni Tutton salt.[12] Similar reduction from the free-ion values was also found for the Slater-Condon parameters for Ni^{2+} in CdS and ZnO from the analysis of the optical spectra.[5] These results point toward the idea of strong bonding between the impurity ion and the ligands. One also notices the fact that the Slater-Condon parameters are appreciably smaller for Ni^{2+} in CdS than in ZnO which is suggestive of a stronger covalency in the former. However, the same sort of difference in λ (-170 ± 10 cm^{-1} versus -175 ± 25 cm^{-1}) is not so obvious from the magnetic measurements, although a larger covalency in Ni^{2+}:CdS is compatible with our results within their limits of accuracy. It is interesting to note that a substantially larger reduction of the spin-orbit constant has been reported in Cu^{2+}:ZnO.[13]

Our analysis leads to the conclusion that the magnetic susceptibilities of Ni^{2+} in CdS and ZnO can be explained quantitatively by using the ligand field theory. The considerable reduction of the spin-orbit coupling constant indicates delocalization of the valence electrons of the Ni^{2+}. It has been pointed out that even in the case of Cu^{2+}:ZnO where an unusually strong bonding between the impurity ions and the ligands is found, the delocalized electrons in Cu^{2+}:ZnO extend only over the Cu^{2+} ion and its nearest neighbors.[13] Thus the use of the ligand field theory in our case should be justifiable. The effect of the covalent bonding is taken into account by making the spin-orbit constant an adjustable parameter. In fact, if one assumes different degrees of covalency between the one-electron orbitals t_2 and e as was done for Cu^{2+}:ZnO, the spin-orbit interaction is then described by two parameters, i.e., $(t_2|H_{s\text{-}0}|t_2)$ and $(t_2|H_{s\text{-}0}|e)$. However, since at $T<500$°K the magnetic susceptibility is governed primarily by the ground state and the two trigonal components of $T_1[^3T_1(F)]$, and since all these two states belong to the $(t_2)^2$ configuration (in the notation of the strong-field scheme), the susceptibility should be rather insensitive to $(t_2|H_{s\text{-}0}|e)$ and hence the value of λ determined here corresponds to the coupling constant of the t_2 state. It has also been customary to introduce an additional parameter, the "orbital reduction factor" k, which relates the matrix elements of the orbital angular momentum evaluated by using the ligand field LCAO MO with those evaluated by the pure atomic functions.[13,14] Generally, this factor is not much less than one even when there is a considerable reduction in λ. For instance, in the case of Fe^{2+} in ZnF_2,[15] the spin-orbit constant is about 60% of the free-ion value and k is around 0.95. Even in the very strong bonding case of Cu^{2+}:ZnO where λ is less than 10% of the value of a free Cu^{2+}, the orbital reduction factor is only 0.46.[13] In our theoretical analysis here, this reduction ratio is not used, i.e., it is taken as unity. When the orbital reduction ratio is included in the theory, all the matrix elements of $\mathfrak{u}$ which are relevant in the calculation of the susceptibility, are modified by the same multiplicative constant. The inclusion of k merely alters the theoretical χ by a constant factor and has the same effect as changing the concentration of the nickel ion. Thus the orbital reduction ratio cannot be obtained from our susceptibility data as the precise concentrations of Ni^{2+} in both ZnO and CdS are not known. On the other hand, the crystal field parameters λ and K affect the shape of the curve of χ versus T rather than the absolute magnitude of the susceptibility, hence the values of these parameters which we gave in the preceding paragraphs still stand correct even when the orbital reduction factor is taken into consideration.

ACKNOWLEDGMENTS

The authors are indebted to Dr. T. L. Estle, Dr. M. deWit, and Dr. H. A. Weakliem for supplying us with the crystals and for many stimulating discussions. Thanks are also due E. C. Segraves and C. R. Yarger for their assistance in the magnetic measurements and data analysis, and to C. F. Dorman for his aid in the point-charge calculations.

---

[12] W. Low, Phys. Rev. **109**, 247 (1948); J. H. E. Griffiths and J. Owen, Proc. Roy. Soc. (London) **A213**, 459 (1952).

[13] R. E. Dietz, H. Kamimura, M. D. Sturge, and A. Yariv, Phys. Rev. **132**, 1554 (1963).

[14] W. Low, in *Solid State Physics*, Suppl. 2, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1960), pp. 99, 100.

[15] M. Tinkham, Proc. Roy. Soc. (London) **A236**, 549 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Electronic Specific Heats of Dilute Solid Solutions

H. JONES*
Department of Physics, Oklahoma State University, Stillwater, Oklahoma
(Received 2 January 1964)

The partition function of an electron gas in the presence of scattering centers, such as exist in a dilute solution, is obtained by using the method first introduced by Matsubara. It is shown that the electronic specific heat at low temperatures, in such a case, contains a term arising from virtual electron scattering. If the Fermi limit in the pure metal lies just beyond a sharp peak in the density-of-states curve, this term increases the specific heat. Comparison with the observations by Rayne on the dilute solid solutions of zinc and germanium in copper indicate that the effect discussed is probably the explanation of the observed initial increase in the low-temperature specific heat at small concentrations.

1. INTRODUCTION

THE electronic properties of disordered alloys are difficult to analyze because there are no simple quantum numbers in terms of which the single-electron wave functions can be described. Equilibrium properties, however, are fully determined when the partition function is known, and this can be obtained as the trace of an operator with respect to any complete orthonormal set of total wave functions. For dilute solid solutions the Slater determinants of the Bloch functions of the pure metal form an obvious choice for such a set. In this representation, the wave function of the ground state of the alloy is not a single Slater determinant, as in the Hartree-Fock approximation for a pure metal, but an infinite series of such determinants. The first member of the series corresponds to the ground state of the pure metal, and the other members to various excited states. In wave-vector k space the alloy has not therefore a sharply defined Fermi surface in its ground state but one which is broadened by the effect of virtual electron scattering. There must, of course, always be a precisely defined Fermi limit in terms of the true one-electron states of the alloy but these are not representable in k space. When the solute concentration is a few percent, the broadening of the Fermi surface of the ground state will be large compared to the thermal broadening in the pure metal at low temperatures.

If the density of states in the pure metal has special features lying just below the Fermi limit, these will not affect the low-temperature electronic specific heat of the pure metal, but they will affect that of the dilute alloy if the ground-state broadening, just referred to, is large enough to include these features. This appears to be exactly the situation which exists in copper and its solid solutions. In pure copper it is believed that the Fermi limit lies just beyond the peak of the density-of-states curve. Hence, in dilute solid solutions, the influence of this peak may be felt on the equilibrium properties and, in particular, on the low-temperature electronic specific heat. It will be shown that in such a case it leads to a slight increase in the specific heat even when the electron concentration remains constant. This increase is closely related to the residual electrical resistance and is determined by the matrix elements of the scattering potential.

2. PERTURBATION EXPANSION OF THE PARTITION FUNCTION

The partition function is obtained as a particularly simple example of the method introduced by Matsubara[1] and developed by Thouless.[2] Let H denote the complete Hamiltonian, N the number of electrons in a volume v, $\beta=1/kT$, and $\alpha=\mu\beta$. If μ denotes the partial potential α is the activity. The grand partition function Z is given by the trace of ψ, where

$$\psi=e^{\alpha N-\beta H}, \tag{1}$$

and

$$Z=\mathrm{Tr}\psi=\sum_n\langle n|e^{\alpha N-\beta H}|n\rangle. \tag{2}$$

The wave functions $|n\rangle$, which are used for the evaluation of the trace are Slater determinants of Bloch functions including the spin factor and relating to the pure metal.

To simplify the notation let

$$H-\mu N=H_0+V(\mathbf{r}), \tag{3}$$

where H_0 is the Hamiltonian of the pure metal minus μN, and $V(\mathbf{r})$ is the potential energy of an electron due to the randomly distributed impurity atoms. $V(\mathbf{r})$ is therefore the difference between the potential in the solid solution and that in the pure metal.

The interaction representation is defined by

$$\psi_I=e^{H_0\beta}\psi, \tag{4}$$

and if we write

$$H_1(u)=e^{H_0u}V(\mathbf{r})e^{-H_0u}, \tag{5}$$

or in terms of the creation and annihilation operators, $\tilde{C}_k(u)$, $C_k(u)$,

$$H_1(u)=\sum_{k,l}\langle k|V|l\rangle\tilde{C}_k(u)C_l(u), \tag{6}$$

then

$$\partial\psi_I/\partial\beta=-H_1(\beta)\psi_I, \tag{7}$$

* On leave of absence from Imperial College, London, England.

[1] T. Matsubara, Progr. Theoret. Phys. (Kyoto) 14, 351 (1955).
[2] D. J. Thouless, Phys. Rev. 107, 1162 (1957); *The Quantum Mechanics of Many-Body Systems* (Academic Press Inc., New York, 1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which may be solved by iteration as follows:

$$\psi_I=1-\int_0^\beta du_1 H_1(u_1)+\int_0^\beta du_1\int_0^{u_1} du_2 H_1(u_1)H_1(u_2)$$
$$-\int_0^\beta du_1\int_0^{u_2} du_2\int_0^{u_2} du_3 H_1(u_1)H_1(u_2)H_1(u_3)+\text{etc.}, \tag{8}$$

where

$$\beta>u_1>u_2>u_3\cdots.$$

The above solution not only satisfies (7) but also the boundary condition, $\psi_I=1$, when $\beta=0$.

From (4) and (2) it follows that, with ψ_I as given by (8),

$$\frac{Z}{Z_0}=\frac{\sum_n\langle n|e^{-H_0\beta}\psi_I|n\rangle}{\sum_n\langle n|e^{-H_0\beta}|n\rangle}, \tag{9}$$

where Z_0 is the partition function relating to the pure metal. The terms of the sum (8) can be evaluated in a manner exactly analogous to that used for the determination of the energy of the ground state in the many-body problem.

Writing

$$H_0=\sum_k(\epsilon_k-\mu)C_k^+C_k,$$

it follows from relations similar to (5), and applicable to any operator, that

$$C_k(u)=C_ke^{-(\epsilon_k-\mu)u},$$
$$\tilde{C}_k(u)=C_k^+e^{(\epsilon_k-\mu)u}. \tag{10}$$

An essential feature of Matsubara's method (cf. Ref. 2) lies in the following definition of the normal product which ensures that its trace in (9) vanishes:

$$N[C_k^+C_k]=(1-f_k)C_k^+C_k-f_kC_kC_k^+. \tag{11}$$

For example, the contribution to Z/Z_0 of any normal product occurring in (8) is given by

$$(1/Z_0)\sum\langle n|e^{-H_0\beta}N[C_k^+C_k]|n\rangle$$
$$=(1/Z_0)\sum\langle n|e^{-H_0\beta}\{(1-f_k)C_k^+C_k-f_kC_kC_k^+\}|n\rangle, \tag{12}$$

which is zero, if

$$(1/Z_0)\sum\langle n|e^{-H_0\beta}C_k^+C_k|n\rangle=f_k,$$
and
$$(1/Z_0)\sum\langle n|e^{-H_0\beta}C_kC_k^+|n\rangle=1-f_k. \tag{13}$$

The number f_k, defined by (13), is just the occupation number for the state k, i.e.,

$$f_k=\{1+e^{\beta(\epsilon_k-\mu)}\}^{-1}. \tag{14}$$

The contractions then introduce factors f_k or $1-f_k$ as indicated below. Using the linked-graph theorem, the result to second order can be given in the following form.

$$\ln(Z/Z_0)=\ln(\sum\text{ all graphs})=\sum(\text{linked graphs})$$
$$=-\beta\sum_k V_{kk}f_k+\beta\sum_{k,l}|\langle k|V|l\rangle|^2$$
$$\times[f_k(1-f_l)/(\epsilon_l-\epsilon_k)]+\text{etc.} \tag{15}$$

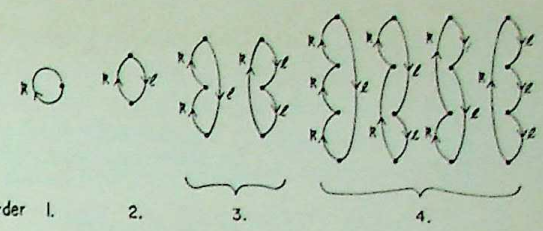


FIG. 1. Diagrams for the perturbation series of the partition function involving only two wave vectors.

The second-order term in (15) diverges logarithmically near the Fermi limit, and it is therefore necessary to make a partial summation over all orders of (8) in which only two wave vectors, k and l, are involved. If we represent the vertices by points with the values of u increasing upwards, the graphs of the terms to be summed are as shown in Fig. 1.

When the operators are u ordered as in (8), it follows from (13) that an upward line corresponds to $C_kC_k^+$ and gives a factor $1-f_k$ and a downward line to $C_l^+C_l$ which gives a factor f_l. The energy denominators for all these particular graphs are powers of $\epsilon_l-\epsilon_k$ and the matrix elements at the vertices either V_{kl}, V_{lk}, V_{kk}, or V_{ll}. [Henceforth we use this abreviated notation rather than that of (15).] It is now easy to see that the above sequence of graphs gives the following contribution to $\ln(Z/Z_0)$.

$$-\beta\sum_k V_{kk}f_k+\beta\sum_{k,l}\frac{|V_{kl}|^2f_k(1-f_l)}{\epsilon_l-\epsilon_k}$$
$$\times\left\{1-\left[\frac{V_{kk}f_k+V_{ll}(1-f_l)}{\epsilon_l-\epsilon_k}\right]+[\ ]^2-[\ ]^3+\cdots\right\}$$
$$=-\beta\sum_k V_{kk}f_k+\beta\sum_{k,l}\frac{|V_{kl}|^2f_k(1-f_l)}{\epsilon_l-\epsilon_k+\{V_{kk}f_k+V_{ll}(1-f_l)\}}, \tag{16}$$

which remains finite at the Fermi limit. It may be noted that the lower limits of the integrals in (8) (i.e., the zeros) give no contribution because factors like $e^{(\epsilon_k-\epsilon_l)\beta}$, etc., combine with f_k and f_l in such a way as to lead to exact cancellation. There are, of course, many other terms in the expansion of ψ_I, but all will involve higher powers of the nondiagonal matrix elements V_{kl}, and will therefore be neglected in the present approximate considerations.

The range of ϵ_k and ϵ_l (now written for convenience as ϵ and η, respectively) over which the summand of the second sum of (16) differs appreciably from zero is shown in Fig. 2.

Since all quantities in (16) apart from V_{kl} depend only on the energies of the states k and l, the first step in the summation is to average $|V_{kl}|^2$ over all relative directions θ of the vectors k and l, and to assume that the result depends only on ϵ and η. This assumption is correct if the matrix elements are calculated with

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

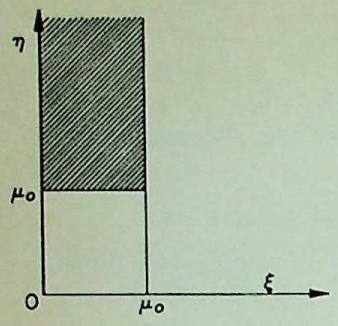


FIG. 2. Range of integration for the second-order correction to the partition function.

respect to free-electron wave functions. It will be assumed, therefore, that

$$\frac{1}{2}\int_0^{\pi} |V_{kl}|^2 \sin\theta d\theta = K(\epsilon,\eta) = K(\eta,\epsilon). \tag{17}$$

To simplify the calculations the following assumptions and approximations are also made: (i) that in the demoninator of (16) f_k and f_l may be replaced by their values at absolute zero, (ii) V_{kk} is independent of k and equal to a constant V. Hence, if $N(\epsilon)$ denotes the density of states, the sum (16) may be written as an integral as follows:

$$\begin{aligned}\int\int \frac{N(\epsilon)N(\eta)K(\epsilon,\eta)}{\eta-\epsilon+2V} f(\epsilon)[1-f(\eta)]d\epsilon d\eta \\ = \int_0^{\mu_0} d\epsilon \int_{\mu_0}^{\infty} d\eta \frac{N(\epsilon)N(\eta)K(\epsilon,\eta)}{\eta-\epsilon+2V} \\ + \int_{\mu_0}^{\infty} d\eta \int_0^{\infty} d\epsilon \frac{N(\epsilon)N(\eta)K(\epsilon,\eta)}{\eta-\epsilon+2V}\phi(\epsilon) \\ - \int_0^{\mu_0} d\epsilon \int_0^{\infty} d\eta \frac{N(\epsilon)N(\eta)K(\epsilon,\eta)}{\eta-\epsilon+2V}\phi(\eta),\end{aligned} \tag{18}$$

where

$$\phi(\epsilon) = \begin{cases} \{1+e^{\beta(\epsilon-\mu)}\}^{-1}, & \epsilon > \mu_0 \\ -\{1+e^{-\beta(\epsilon-\mu)}\}^{-1}, & \epsilon < \mu_0, \end{cases} \tag{19}$$

and the second and third integrals are taken along the boundary of the shaded region of Fig. 2. The term arising from the immediate neighborhood of (μ_0,μ_0) has been omitted as negligible compared with the others. In the last integral the symbols ϵ and η may be interchanged; and therefore, if

$$g(\epsilon) = \int_0^{\mu_0} \frac{N(\eta)K(\epsilon,\eta)}{\epsilon-\eta+2V} d\eta + \int_{\mu_0}^{\infty} \frac{N(\eta)K(\epsilon,\eta)}{\epsilon-\eta-2V} d\eta, \tag{20}$$

the whole of the second sum of (16) may be written

$$-\beta E_2 - \beta \int_0^{\infty} N(\epsilon)g(\epsilon)\phi(\epsilon)d\epsilon, \tag{21}$$

where E_2 is the second-order energy due to the perturbation and is independent of temperature. If E_1 denotes the energy correction in the first order, then

$$\ln Z = \ln Z_0 - \beta(E_1+E_2) - \beta \int_0^{\infty} N(\epsilon)g(\epsilon)\phi(\epsilon)d\epsilon. \tag{22}$$

The last term of (22) can be written

$$\begin{aligned}-\beta \int_0^{\infty} N(\epsilon)g(\epsilon)\phi(\epsilon)d\epsilon \\ = \beta \int_0^{\mu} \frac{N(\epsilon)g(\epsilon)d\epsilon}{1+e^{-\beta(\epsilon-\mu)}} - \beta \int_{\mu}^{\infty} \frac{N(\epsilon)g(\epsilon)d\epsilon}{1+e^{\beta(\epsilon-\mu)}}.\end{aligned} \tag{23}$$

Let $x=\beta(\epsilon-\mu)$, then

$$\begin{aligned}\beta \int_0^{\infty} N(\epsilon)g(\epsilon)\phi(\epsilon)d\epsilon = \int_0^{\infty} \frac{dx}{1+e^x}\left[N\left(\mu+\frac{x}{\beta}\right)g\left(\mu+\frac{x}{\beta}\right)\right. \\ \left. - N\left(\mu-\frac{x}{\beta}\right)g\left(\mu-\frac{x}{\beta}\right)\right],\end{aligned} \tag{24}$$

in which the usual approximation has been made of replacing the lower limit $-\beta\mu$ by $-\infty$ in one integral.

The partition function for the pure metal is well known and may be expressed

$$\begin{aligned}\ln Z_0 = \beta\{\mu n(\mu) - E(\mu)\} + \frac{1}{\beta}\int_0^{\infty} \ln(1+e^{-x}) \\ \times\left[N\left(\mu+\frac{x}{\beta}\right)+N\left(\mu-\frac{x}{\beta}\right)\right]dx,\end{aligned} \tag{25}$$

where

$$n(\mu) = \int_0^{\mu} N(\epsilon)d\epsilon, \quad \text{and} \quad E(\mu) = \int_0^{\mu} N(\epsilon)\epsilon d\epsilon, \tag{26}$$

are both functions of temperature through their dependence on μ. On substituting (24) and (25) into (22) it follows that for low temperatures,

$$\begin{aligned}\ln Z = \alpha n(\mu) - \beta[E(\mu)+E_1+E_2] \\ + \frac{\pi^2}{6\beta}\left\{N(\mu) - \left[\frac{d}{d\epsilon}N(\epsilon)g(\epsilon)\right]_{\mu}\right\}.\end{aligned} \tag{27}$$

It may be noted that $N(\mu)$ is the density of states including the factor 2 which arises as a consequence of each state being doubly occupied by electrons of opposite spins.

The properties of an assembly consisting of a fixed number of electrons are obtained by using the relation,

$$(\partial \ln Z/\partial\alpha)_{\beta} = \bar{n}, \tag{28}$$

to determine μ when $\bar{n}$ is constant. The energy is then given by

$$-(\partial \ln Z/\partial\beta)_{\alpha} = \bar{E}. \tag{29}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The specific heat of the system C_v (it has been assumed throughout that the volume is constant) is given by

$$C_v=-\frac{\beta^2}{k}\left(\frac{\partial\bar{E}}{\partial\beta}\right)_{\bar{n}}=\gamma T. \tag{30}$$

Although γ may readily be evaluated, it is sufficient for the present purpose to note that

$$\gamma=\gamma_0\left\{1-\frac{1}{N(\mu)}\left[\frac{d}{d\epsilon}N(\epsilon)g(\epsilon)\right]_\mu\right\}, \tag{31}$$

where γ_0 refers to the pure metal.

3. EVALUATION OF MATRIX ELEMENTS

It will be assumed that the potential $V(\mathbf{r})$ arises from ν screened unit positive charges randomly distributed through the volume v which contains N electrons and the same number of lattice sites. The solute concentration is therefore given by $c=\nu/N$. Hence

$$V(\mathbf{r})=-e^2\textstyle\sum_i \exp\{-\alpha|\mathbf{r}-\mathbf{R}_i|\}/|\mathbf{r}-\mathbf{R}_i|, \tag{32}$$

where $\mathbf{R}_i$ denotes the position of the solute atoms, and α is the reciprocal of the screening length.

If in order to estimate the magnitude of V_{kl} we use free-electron wave functions defined by

$$\psi_{\mathbf{k}}=\frac{1}{\sqrt{v}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{33}$$

the following well-known result is readily obtained.

$$V_{kl}=-\frac{4\pi e^2}{v}\frac{1}{|\mathbf{k}-\mathbf{l}|^2+\alpha^2}\textstyle\sum_i e^{i(\mathbf{k}-\mathbf{l})\cdot\mathbf{R}_i}. \tag{34}$$

As a result of the random nature of the distribution of solute atoms,

$$|V_{kl}|^2=\left(\frac{4\pi e^2}{v}\right)^2\frac{\nu}{(|\mathbf{k}-\mathbf{l}|^2+\alpha^2)^2}. \tag{35}$$

The average of this expression over all directions of $\mathbf{l}$ relative to $\mathbf{k}$ depends only on k^2 and l^2 and therefore is in accordance with assumption (17).

In order to simulate the effect of the peak in the density of states of copper just below the Fermi limit and at the same time enable the integrals to be evaluated without too much difficulty we adopt the following model.

$$N(\eta)=N_0+Np\delta(\eta-\epsilon_1), \quad \epsilon_1<\mu_0, \tag{36}$$

where N_0 is a constant. Here p is a fraction of order 10^{-1}. It is further assumed that $K(\mu,\eta)\to 0$ for large values of η as would be the case with (35). Substituting (36) into (20) gives

$$g(\epsilon)=\int_0^{\mu_0}\frac{N_0K(\epsilon,\eta)}{\epsilon-\eta+2V}d\eta+\int_{\mu_0}^{\infty}\frac{N_0K(\epsilon,\eta)}{\epsilon-\eta-2V}d\eta+\frac{NpK(\epsilon,\epsilon_1)}{\epsilon-\epsilon_1+2V}. \tag{37}$$

When $\epsilon\simeq\mu_0$ it will be seen that the first two integrals are of opposite sign and tend to cancel each other, and their contribution to $g'(\mu)$ will not be considered further. The third term represents the effect of the peak in $N(\epsilon)$ and its variation with ϵ results mainly from the denominator. Hence a reasonable approximation is

$$g'(\mu)=-NpK(\mu,\epsilon_1)/(\mu-\epsilon_1+2V)^2, \text{ and } N'(\mu)=0. \tag{38}$$

Thus

$$\gamma=\gamma_0\left\{1+\frac{K(\mu,\epsilon_1)Np}{(\mu-\epsilon_1+2V)^2}\right\}. \tag{39}$$

To estimate the additional term in γ we use (35) and (17) and find

$$\frac{K(\mu,\mu)Np}{\mu^2}=\frac{(4\pi e^2)^2}{\Omega^2}\left(\frac{\hbar}{2m}\right)^2\frac{p}{\mu^2(w^2+4w\mu)}\left(\frac{\nu}{N}\right), \tag{40}$$

where Ω is the atomic volume which may be written $\Omega=\frac{4}{3}\pi r_s^3$, and $w=(\hbar^2/2m)\alpha^2$. Hence using the free-electron approximation to determine μ_0 and writing for simplicity, $w=\mu$,

$$\frac{K(\mu,\mu)Np}{\mu^2}=\frac{p}{5}\left(\frac{256}{4\pi^4}\right)^{1/3}\frac{(e^2/r_s)^2}{\mu^2}c, \tag{41}$$

which gives finally, neglecting $2V$ compared with $\mu-\epsilon_1$,

$$\gamma=\gamma_0\left\{1+\frac{0.13p}{(1-\epsilon_1/\mu)^2}\cdot\frac{(e^2/r_s)^2}{\mu^2}c\right\}. \tag{42}$$

Suppose the ratio of the two energies is unity, that ϵ_1 lies $\frac{1}{10}$ of μ_0 below the Fermi limit, and that $p\simeq 10^{-1}$, i.e., one-tenth of the electrons lie under the effective part of the peak. These rough estimates which seem reasonable enough then give $\gamma=\gamma_0(1+c)$, where c is the concentration expressed as a fraction.

4. RELATION TO RESIDUAL RESISTIVITY

Imagine a hole to exist in the Fermi distribution at absolute zero in the neighborhood of ϵ_1. Elastic scattering due to impurities will result in a finite lifetime τ of any particular state k of this hole. This lifetime will be approximately equal to the relaxation time for impurity resistivity. Each level therefore will have a width of order $\hbar/\tau$. The delta function of (36) is therefore effectively replaced by a function of width $\hbar/\tau$, and if $\mu-\epsilon$, is small part of the tail will overlap the Fermi limit and give a contribution to the linear electronic specific heat. This argument can be made semiquantitative as follows. The linewidth at ϵ may be written

$$\frac{\hbar/\tau}{\pi\{(\epsilon-\eta)^2+(\hbar/\tau)^2\}}, \tag{43}$$

and hence if the density of states in the pure metal is

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$N_0(\eta)$ in the solid solution it is

$$N(\epsilon)=\int_0^\infty \frac{N_0(\eta)\hbar/\tau}{\pi\{(\epsilon-\eta)^2+(\hbar/\tau)^2\}}d\eta. \tag{44}$$

For the model in which the density-of-states $N_0(\eta)$ is as given by (36) it follows that

$$N(\mu)=N_0+\frac{(\hbar/\tau)Np}{\pi\{(\mu-\epsilon_1)^2+(\hbar/\tau)^2\}}. \tag{45}$$

Hence when $\hbar/\tau \ll \mu-\epsilon_1$, which implies that the correction is small and corresponds to a second-order effect,

$$N(\mu)=N_0+\frac{Np(\hbar/\tau)}{\pi\mu^2(1-\epsilon_1/\mu)^2} \tag{46}$$

and

$$\gamma=\gamma_0\left\{1+\frac{p(\hbar/\tau)}{\pi\mu(1-\epsilon_1/\mu)^2}\right\}, \tag{47}$$

where the approximation $N_0\mu \simeq N$ has been used. The relaxation time is given by the well-known formula,

$$\frac{\hbar}{\tau}=\frac{mkv}{2\pi\hbar^2}\int_0^\pi |V_{kl}|^2(1-\cos\theta)\sin\theta d\theta, \tag{48}$$

where $v=\Omega N$ is the volume which contains ν impurities and with respect to which the electronic wave functions are normalized. If we define half of the integral as $K'(\mu,\mu)$, then a simple calculation shows that

$$\hbar/\tau=3\pi NK'/2\mu, \tag{49}$$

and

$$\gamma=\gamma_0\{1+[3K'Np/2(\mu-\epsilon_1)^2]\}, \tag{50}$$

which is similar to (39) apart from the numerical factor 3/2, the neglect of $2V$ in the denominator and the different definitions of K and K'. Both derivations show that the addition to the electronic specific heat is proportional to the residual resistivity.

5. COMPARISON WITH EXPERIMENTAL RESULTS

The electronic specific heats of dilute solid solutions of Zn and Ge in Cu have been measured by Rayne[3] and by Veal and Rayne.[4] It is found that initially for small concentrations of the solute γ increases in both cases. Since the Fermi limit lies beyond the peak of the density of states a simple theory would suggest a decrease of γ due to the increase of the electron concentration. This decrease is easily calculated and is indicated by Veal and Rayne. If we attribute the difference between the observed increase and the expected decrease to virtual electron scattering and express the results for the initial slopes by giving the values of b in $\gamma/\gamma_0=1+bc$, where c is the fractional concentration, then for Zn we find $b=1.4$ and for Ge, $b=5.8$. Equation (42) appears therefore to give values of b of roughly the right magnitude. Moreover, as might be expected, Ge with the greater scattering power has the higher value.

There are other examples where the observed electronic specific heat is larger than would be anticipated from a simple band model. For example, Keesom and Kurrelmeyer[5] find for a copper-nickel alloy containing approximately 20 at.% Ni, a γ value which is 1.56 times as great as the γ value for pure Cu. Conventional theory visualizes the d band as just filled in $Cu_{60}Ni_{40}$. At 80% Cu the Fermi limit will be beyond the point where $N(\epsilon)$ rises steeply due to the d levels. A similar situation to that discussed in the preceding theory therefore exists and may perhaps be explained in the same way. The measurements of Hoare and his collaborators[6,7] on Ag-Pd and Au-Pt point in the same direction.

ACKNOWLEDGMENTS

The author wishes to thank the members of the Physics Department of the Oklahoma State University for their kind hospitality and the National Science Foundation for the award of a Senior Foreign Scientist Fellowship.

[3] J. A. Rayne, Phys. Rev. **108**, 22 (1957).
[4] B. W. Veal and J. A. Rayne, Phys. Rev. **130**, 2156 (1963).
[5] W. K. Keesom and B. Kurrelmeyer, Physica **7**, 1003 (1940).
[6] F. E. Hoare and B. Yates, Proc. Roy. Soc. (London) **A240**, 42 (1957).
[7] D. W. Budworth, F. E. Hoare, and J. Preston, Proc. Roy. Soc. (London) **A257**, 250 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Motional Narrowing of Nuclear Magnetic-Resonance Lines in Manganese-Doped Lithium Fluoride*

T. G. Stoebe, T. O. Ogurtani, and R. A. Huggins
Department of Materials Science, Stanford University, Stanford, California
(Received 6 January 1964)

Nuclear magnetic-resonance experiments have been performed between 25 and 550°C in LiF powder doped with manganese. Motional narrowing of the Li^7 resonance line occurred somewhat below 400°C. Resonance absorption lines were not observable in the temperature range 200–390°C, probably because an effective loss of paramagnetic relaxation centers increased the spin-lattice relaxation time. Electron spin-resonance experiments verified that the paramagnetic manganese does cluster upon heating above 200°C. A method of analysis is developed which eliminates errors due to constant linewidth contributions in the analysis of motional narrowing data in terms of diffusion activation energies. The activation energy for motion of Li ion vacancies in LiF is found to be 0.74 eV.

I. INTRODUCTION

MOTIONAL narrowing of nuclear magnetic-resonance lines has been used to investigate extrinsic diffusion in several doped ionic crystals. Examples include[1] LiBr where vacancies present on Li sites gave rise to extrinsic Li-ion diffusion, and[2,3] CaF_2 where F ions were seen to diffuse. In the present study, motional narrowing is observed in manganese-doped LiF powder. The substitution of Mn^{2+} ions for Li^+ ions in LiF gives rise to an equal number of Li-ion vacancies to maintain charge neutrality. For this reason, it is possible to interpret motional narrowing in this system in terms of the diffusion of Li-ion vacancies.

II. EXPERIMENTAL TECHNIQUES

Powdered LiF samples were prepared by precipitation with NH_4F from an aqueous solution of LiCl containing 0.1% $MnCl_2$. They were subsequently filtered, washed, and then dried at 100°C for a short time. The Li^7 nuclear magnetic-resonance absorption line was observed at 7.5 Mc/sec using a modified Varian wide-line spectrometer. Experiments were performed between 25 and 550°C by inserting an air-cooled high-temperature Dewar in the Varian probe.

III. METHOD OF DATA ANALYSIS

To date, the analysis of motional narrowing curves to give diffusion coefficients has been only moderately successful. In metallic aluminum[4] and lithium,[5,6] experimental activation energies for diffusion have varied from values obtained by other techniques[7–9] by up to 30%. This is due both to changes in lineshape and constant contributions to the linewidth, which make data analysis uncertain.

A method has been developed which enables analysis of motional narrowing data to produce accurate values of diffusion activation energies. It eliminates errors due to any constant contributions to the linewidth, including magnetic-field inhomogeneities, and is applicable in temperature regions where the lineshape is nearly constant.

Since narrowed lines are nearly Lorentzian,[10] the actual dipolar linewidth δH_d can be related to the diffusion coefficient D by the relation

$$\delta H_d = C\delta H_0 D^{-1} = A\exp(E/kT), \tag{1}$$

where E is the activation energy for motion of Li-ion vacancies. The constant A contains the concentration of vacancies in the system and the factor relating the linewidth to $(1/T_2)$. The relation between the measured line width δH_m and the actual dipolar linewidth can be written

$$\delta H_m = a\delta H_d + b\delta H_1, \tag{2}$$

where δH_1 is a constant.

One of the methods of data analysis that has been used by others consists of plotting $\ln\delta H_m$ versus $1/T$ and equating the slope of the curve to E/k. If, however, a constant addition to the line width is present such as given in (2), the slope S obtained in this manner is not E/k but rather

$$S = \frac{d\ln\delta H_m}{d(1/T)} = \frac{E}{k}\frac{1}{[1+(b\delta H_1/aA)e^{-E/kT}]}. \tag{3}$$

Equation (3) shows that S becomes smaller than E/k at higher temperatures, which is the region of interest in the present case.

* This work was partially supported by the U. S. Atomic Energy Commission as well as the Advanced Research Projects Agency, through the Center for Materials Research at Stanford University.

[1] R. R. Allen and M. J. Weber, J. Chem. Phys. 38, 2970 (1963).
[2] W. J. Veigele and A. W. Bevan, Jr., Phys. Rev. 131, 1585 (1963).
[3] W. J. Veigele, Bull. Am. Phys. Soc. 9, 25 (1964).
[4] E. F. W. Seymour, Proc. Phys. Soc. (London) A66, 85 (1953).
[5] H. J. Sukenik, thesis, University of California, 1958 (unpublished).
[6] H. S. Gutowsky and B. R. McGarvey, J. Chem. Phys. 20, 1472 (1952).
[7] C. P. Flynn and E. F. W. Seymour, Proc. Phys. Soc. (London) A77, 922 (1961).
[8] J. J. Spokas and C. P. Slichter, Phys. Rev. 113, 1462 (1959).
[9] D. F. Holcomb and R. E. Norberg, Phys. Rev. 98, 1074 (1955).
[10] R. Kubo and K. Tomita, J. Phys. Soc. Japan 9, 888 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

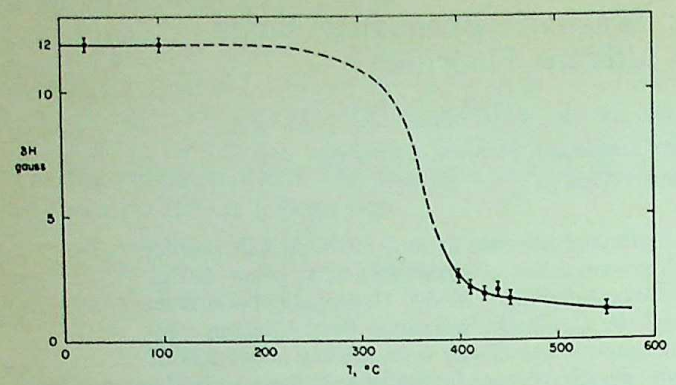


FIG. 1. Linewidth of Li^7 resonance line in Mn-doped LiF as a function of temperature.

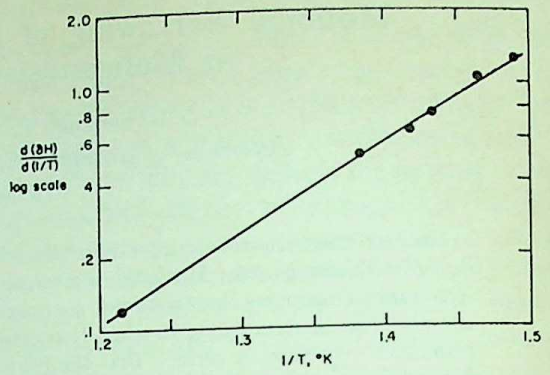


FIG. 2. Plot of $\ln[d\delta H/d(1/T)]$ versus $1/T$ for the data of Fig. 1, which yields activation energy of 0.74 eV. (*Note added in proof.* The scale on abscissa and ordinate should contain a factor of 10^3.)

If, however, δH_m is first differentiated with respect to $1/T$, followed by taking the logarithm of this derivative, we have

$$\ln\Delta \equiv \ln\frac{d\delta H_m}{d(1/T)} = \text{const} + \frac{E}{kT}. \qquad (4)$$

The slope S of a plot of $\ln\Delta$ versus $1/T$ will now yield the true value of E/k, independent[11] of δH_1. This analysis is easily performed, since Δ can be evaluated at several temperatures on a plot of the experimental data for δH_m versus $1/T$.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The temperature variation of the observed linewidth, defined as the width in gauss between the points of maximum and minimum slope, is given in Fig. 1. The limits of error shown were determined from estimates of the experimental accuracy in determining the linewidth and from the variations between values obtained from several experimental measurements at the same temperature. The room-temperature values of linewidth may be compared with the half max-min linewidth of about 6 G given by Chang *et al.*[12,13] in similar samples at 77°K. Nuclear-resonance lines of F^{19} in this material narrowed by about 25% in the same temperature range in which the Li^7 narrowed.

Absorption lines could not be obtained in the temperature interval 200–390°C because of a long spin-lattice relaxation time. This was probably caused by an effective loss of paramagnetic-relaxation centers from the system. Electron spin-resonance experiments were performed, indicating that the manganese ions are uniformly distributed in the lattice at room temperature. After heating the samples for short times above 200°C, the electron spin-resonance line observed at room temperature loses its hyperfine structure due to a strong exchange interaction. This is probably due to clustering of Mn in the Li sublattice, similar to the clustering observed by Jennings[14] in Mn-doped LiCl single crystals. The electron spin-resonance spectra before heating are very similar to those observed by Chang *et al.*[12,13]

The analysis described in Sec. III has been performed on the high-temperature data of Fig. 1. The resulting straight line, shown in Fig. 2, yields an activation energy of 0.74 eV, with a precision of about 10%. This is in good agreement with the activation energies obtained in the extrinsic region by investigators using other techniques; electrical-conductivity results[15] have given $E=0.65$ eV and spin-echo experiments in single crystals[16] have given 0.71 eV. By comparison of the present data with the results of electrical conductivity experiments,[15] it can be determined that the vacancy concentration in the present work is about 5×10^{-5}.

[11] It should be noted that if we include the bilinear term in the expansion of Eq. (2), $(c\delta H_d\delta H_1)$, nothing is changed in the result except the constant a.

[12] T. T. Chang, W. H. Tanttila, and J. S. Wells, J. Chem. Phys. **39**, 2453 (1963).

[13] T. T. Chang, thesis, University of Colorado, 1962 (unpublished).

[14] D. A. Jennings, thesis, University of Colorado, 1962 (unpublished).

[15] Y. Haven, Rec. Trav. Chim. **69**, 1471 (1950).

[16] M. Eisenstadt, Phys. Rev. **132**, 630 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Localized Mode in an Anharmonic Crystal*†

WILLIAM M. VISSCHER
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico
(Received 8 April 1963, revised manuscript received 17 January 1964)

We consider a harmonic crystal with a single isotopic impurity whose coupling to its neighbors includes a cubic anharmonic term. The impurity is chosen to be light enough that a localized mode exists. The Hamiltonian for this system is truncated, in a physically reasonable way, so as to yield an exactly solvable Schrödinger equation. The solutions give a new insight into the nature of the localized mode, which in this approximation is exactly analogous to an unstable particle in the Lee model of field theory. The wave functions are exhibited and the lifetime is calculated for a light impurity in a simple cubic lattice at $T=0$.

I. INTRODUCTION

AN isolated impurity in a harmonic crystal gives rise to certain changes in the vibration spectrum[1–5] of the system. In particular, if the impurity is lighter than the host-crystal atoms and the impurity-mass to host-mass ratio is less than a certain critical number dependent on the atomic arrangement, then a so-called localized mode appears above the quasicontinuum of phonons. The characteristics of this mode are that when it is excited only the atoms near the impurity participate in the motion (the attenuation length decreases as the frequency increases); and that its frequency is higher than the maximum phonon frequency. The latter may be looked upon as the reason why the mode *is* localized; its frequency is too high to propagate in the lattice. The localized mode is a single normal mode of the system, with a sharp frequency.

It is the purpose of this note to investigate what happens to the localized mode when the interatomic force is allowed to contain certain anharmonic terms.[6] The "single-particle" description of independent oscillators (harmonic forces only) then goes over into the analog of an interacting many-body system, and the physics and mathematics encountered is similar to that which occurs in a wide variety of problems. The localized mode state of the harmonic approximation becomes mixed with the multiphonon states whose energies lie nearby. The strength of the anharmonic interaction determines the width of the energy interval in which the mixing occurs; the energy width in turn governs the coherence of the localized mode (now only an approximate eigenstate) and therefore its lifetime.

Lowest order perturbation theory involves the calculation of the diagram shown in Fig. 1, with time increasing from left to right. The single localized mode which is present in the initial state (dashed line) decays into two phonons in the final state (wavy lines) through a cubic anharmonic interaction. The matrix element of the interaction operator

$$h=\sum_{lm} B_{klm}a_m^\dagger a_l^\dagger a_k \tag{1}$$

between a localized mode state (k) and a state in which two phonons (l and m) are present is all that is involved in the perturbation calculation. Such a calculation has been performed by Klemens,[6] who used the phonon operators of the perfect crystal, and the localized mode operator[7] of Montroll and Potts.[2]

The interaction Eq. (1) is one of the terms which arise from the expansion of a cubic anharmonic term in the lattice potential energy

$$H'=\sum_{ij} C_{ij}(\mathbf{x}_i-\mathbf{x}_j)^3 \tag{2}$$

in terms of the phonon and localized mode creation and annihilation operators. The cubic anharmonic term has no lower bound for large relative displacements; hence one would not expect any solutions of a problem in which it is the only perturbation. For our anharmonic interaction, however, we choose only a part of Eq. (2),

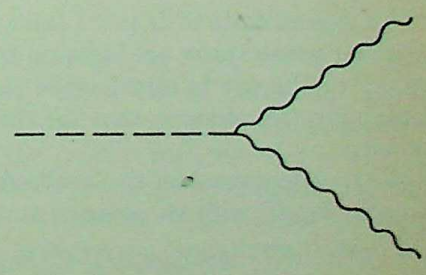


FIG. 1. Feynman diagram for decay of localized mode into two phonons.

* Work done under the auspices of the U. S. Atomic Energy Commission.

† This work was reported at the Moscow Conference on Solid State Theory, December 2–12, 1963.

[1] I. M. Lifshitz, Nuovo Cimento, Suppl. 3, 716 (1956).

[2] E. W. Montroll and R. B. Potts, Phys. Rev. 100, 525 (1955).

[3] Yu. Kagan and Ya. A. Iosilevskii, Zh. Eksperim. i Teor. Fiz. 42, 259 (1962) [English transl.: Soviet Phys.—JETP 15, 182 (1962)].

[4] R. Brout and W. M. Visscher, Phys. Rev. Letters 9, 54 (1962).

[5] Yu. Kagan and Ya. A. Iosilevskii, Zh. Eksperim. i Teor. Fiz. 44, 284 (1963) [English transl.: Soviet Phys.—JETP 17, 195 (1963)].

[6] This question has been studied before by different methods. See P. G. Klemens, Phys. Rev. 122, 443 (1961) and A. A. Maradudin, *1962 Brandeis Lectures in Theoretical Physics* (W. A. Benjamin and Company, Inc., New York, 1963), Vol. 2.

[7] The localized mode displacements assumed by Klemens were those given by Montroll and Potts as asymptotic forms, except that he inadvertently omitted the alternation of sign with each successive lattice distance from the impurity. This omission tends to make Klemens' result for the width too large, as does the fact that he used phonon operators for the perfect harmonic crystal rather than those for the harmonic crystal with a single isolated impurity.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

such that the resulting systems may easily be solved exactly. Our method exhibits the solutions for the wave function explicitly, hence has an esthetic advantage over a perturbation calculation, although the quantitative results should be taken only to lowest order in the anharmonicity parameters.

We shall keep only Eq. (1) and its Hermitian conjugate and consider their sum to be the anharmonic interaction and drop the rest of the terms in Eq. (2). The principal justification for this is that it allows an exact solution to be easily obtained, but it is possible to argue as follows that at least some of the dropped terms are unimportant. The anharmonic interaction is large only when the displacements are large, and the displacement of any atom in the crystal due to the excitation of a single phonon is proportional to $N^{-1/2}$, where N is the total number of atoms. On the other hand, the localized mode displacements do not vanish as one lets $N\to\infty$, but approach a finite limit which decreases more or less exponentially with distance from the impurity atom. It might therefore be reasonable to expect that the most important terms in the expansion of Eq. (2) are those which involve the creation or annihilation of one or more localized modes. Most of the terms in the expansion of Eq. (2) do not satisfy this criterion.

We therefore take

$$H'=h+h^\dagger \tag{3}$$

as our anharmonic interaction. The basic Feynman diagram is Fig. 1 and the time-reverse of it, and we have essentially defined a Lee model[8,9] of the anharmonic crystal.

The total Hamiltonian for this model is

$$H=H_0+H', \tag{4}$$

where H_0 is the harmonic Hamiltonian

$$H_0=\sum_s \hbar\omega_s(n_s+\tfrac{1}{2}),$$

$$n_s=a_s^\dagger a_s.$$

II. EIGENSTATES AND EIGENVALUES

The ground state of H is the same as that of H_0, and the one-phonon states are identical too. It is only when we get to the one localized mode state and the states containing two phonons that the effects of H' begin to be felt.

We therefore consider the localized mode state, which is mixed by H' with two phonon states, and let

$$\Psi=C_k\Phi_k+\sum C_{lm}\Phi_{lm}, \tag{5}$$

where the C's are constants subject to a normalization condition

$$|C_k|^2+\sum_{lm}|C_{lm}|^2=1, \tag{6}$$

and

$$\Phi_k=a_k^\dagger\Phi_0,$$
$$\Phi_{lm}=2^{-1/2}a_l^\dagger a_m^\dagger\Phi_0, \tag{7}$$

with Φ_0 the ground state, and k always referring to the localized mode, l, m to phonons in the quasicontinuum.

By requiring that Eq. (5) be an eigenstate of H with eigenvalue E one immediately obtains

$$(E-\epsilon_k)C_k=2^{1/2}\sum_{lm}B_{klm}^\dagger C_{lm},$$
$$(E-\epsilon_{lm})C_{lm}=2^{1/2}B_{klm}C_k, \tag{8}$$

where ϵ_k and ϵ_{lm} are eigenvalues of H_0 with Φ_k and Φ_{lm}, respectively. Therefore unless E happens to coincide with one of the ϵ_{lm}'s or C_k vanishes we must have

$$E-\epsilon_k=2\sum_{lm}\frac{|B_{klm}|^2}{E-\epsilon_{lm}}, \tag{9}$$

which has as many solutions $E=E_\lambda$ as there are terms in the sum, plus one. Equation (9) is very similar to that obtained by Glaser and Källen[9] in their investigation of a Lee model[8] for an unstable particle. A graph of the right- and left-hand sides of Eq. (9) appears in Fig. 2. The intersections define the set of eigenvalues E_λ.

Certain features of Fig. 2 deserve comment. One is that for a given size crystal it is always possible to have a single negative eigenvalue E_λ for the anharmonic coefficients B_{klm} large enough. However, because as it is defined in Eq. (1) B_{klm} should decrease as $N^{-1/2}$ in order to correspond with Eq. (2) with C_{ij} independent of N; the point of appearance of the negative eigenvalue as a function of anharmonic-force strength will be independent of crystal size for large crystals. The negative eigenvalue appears only for very large B_{klm}. Its existence is due to the fact that our Hamiltonian is not positive definite. Another feature is that an eigenvalue pops up out of the quasicontinuum, reminiscent of the appearance of the localized mode in the harmonic case.

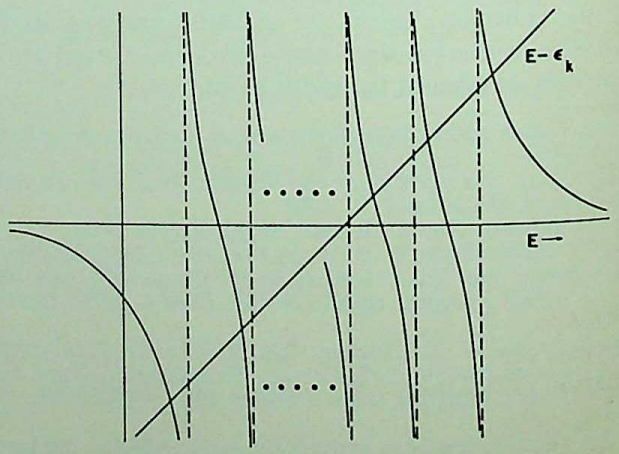


FIG. 2. Graphical solution of Eq. (9). The lowest asymptote will generally be at an energy of order $\hbar\omega_L/N$, the highest slightly below $2\hbar\omega_L$, where ω_L is the maximum phonon frequency.

[8] T. D. Lee, Phys. Rev. 95, 1329 (1954).
[9] V. Glaser and G. Källen, Nucl. Phys. 2, 706 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

This state appears at a smaller, but still unphysically large, value of the anharmonic coupling. Like the negative eigenvalue, the critical-coupling parameter at which the positive energy state appears becomes independent of crystal size for large crystals.

III. COMPLETENESS

Since it is our purpose to study the evolution of the localized mode state as anharmonic forces are introduced, it behooves us to be sure that the set of states given by Eqs. (5) and (8) really spans that part of the Hilbert space which can be reached by operating repeatedly with the anharmonic Hamiltonian on the harmonic localized mode state Φ_k. To do this it is sufficient to show that

$$\sum_\lambda \Psi_\lambda \Psi_\lambda^\dagger = \Phi_k \Phi_k^\dagger + \sum_{lm} \Phi_{lm} \Phi_{lm}^\dagger , \tag{10}$$

where Ψ_λ is an eigenfunction of the Hamiltonian [Eq. (4)] with eigenvalue E_λ [a root of Eq. (9)], and which has the form of Eq. (5). The calculation is straightforward and proceeds as follows. Upon substitution of Eq. (5) into (10), we obtain

$$\begin{aligned}\sum_\lambda \Psi_\lambda \Psi_\lambda^\dagger = &\sum_\lambda C_k(\lambda) C_k^\dagger(\lambda) \Phi_k \Phi_k^\dagger \\ &+ \sum_{lm} C_{lm}(\lambda) C_{lm}^\dagger(\lambda) \Phi_{lm} \Phi_{lm}^\dagger \\ &+ \sum_{l'm' \neq lm} C_{l'm'}(\lambda) C_{lm}^\dagger(\lambda) \Phi_{l'm'} \Phi_{lm}^\dagger \\ &+ \sum_{lm} C_k(\lambda) C_{lm}^\dagger(\lambda) \Phi_k \Phi_{lm}^\dagger \\ &+ \sum_{lm} C_{lm}(\lambda) C_k^\dagger(\lambda) \Phi_{lm} \Phi_k^\dagger , \end{aligned} \tag{11}$$

and illustrate by calculating the coefficient of $\Phi_k \Phi_k^\dagger$.

We introduce a function $C_k(E)$ of the complex variable E, which has the property that it coincides with $C_k(\lambda)$ at $E=E_\lambda$, and is analytic in a small region surrounding each E_λ. Then

$$\sum_\lambda C_k(\lambda) C_k^\dagger(\lambda) = \frac{1}{2\pi i} \int_C C_k(E) C_k^\dagger(E) d \ln D(E) , \tag{12}$$

where

$$D(E) = E - \epsilon_k - F(E) , \tag{13}$$

and

$$F(E) = 2 \sum_{lm} \frac{|B_{klm}|^2}{E - \epsilon_{lm}} , \tag{14}$$

The contour C encircles the roots of $D(E)$ in a counterclockwise sense, and excludes any singularities of $C_k(E)$. Equation (12) follows from the fact that the eigenvalues E_λ are the roots of $D(E)$. Now, from Eqs. (6) and (8) we see that to within a phase factor

$$C_k(E) = [1 - F'(E)]^{-1/2} \tag{15}$$

has the properties we require for $C_k(E)$, and therefore

$$\sum_\lambda C_k(\lambda) C_k^\dagger(\lambda) = \frac{1}{2\pi i} \int_C \frac{dE}{E - \epsilon_k - F(E)} = 1 , \tag{16}$$

which follows by expanding the contour to an infinite circle. Using some care in excluding the singularities of the C's, the other coefficients in Eq. (11) can be similarly shown to be unity or to vanish, and Eq. (10) follows.

IV. LIFETIME OF LOCALIZED MODE

If we imagine that in the anharmonic crystal we can somehow excite the *harmonic* localized mode, its lifetime may easily be calculated. To do this it suffices to calculate

$$(\Psi(0), \Psi(t)) = (\Phi_k, \Psi(t)) , \tag{17}$$

with

$$\Psi(t) = \sum_\lambda C(\lambda) \Psi_\lambda e^{-iE_\lambda t} , \tag{18}$$

and the C's chosen so that

$$\Psi(0) = \Phi_k . \tag{19}$$

Equation (19) implies that

$$C(\lambda) = C_k^\dagger(\lambda) , \tag{20}$$

and the scalar product of Eq. (17) is easily seen to be

$$(\Psi(0), \Psi(t)) = \frac{1}{2\pi i} \int_C e^{-iEt} \frac{1}{E - \epsilon_k - F(E)} dE . \tag{21}$$

Because of the presence of the exponential we can't evaluate the integral as easily as before, and must carefully study the denominator in the integrand.

If we are interested in times short compared to the longest periods of vibration of the crystal, then we may as well consider the infinite crystal. Therefore, we study the behavior of $F(E)$ in the limit $N \to \infty$. It will become an integral

$$F(E) = \int_0^{2\hbar\omega_L} \frac{\rho(k,\epsilon) d\epsilon}{E - \epsilon} \tag{22}$$

with $\rho(k,\epsilon)$ real and positive, of form similar to that encountered in many fields, but in particular in the theory of lattice-vibration spectra. $F(E)$ has acquired a cut in the locus of the poles it had before $N \to \infty$; the poles, in other words, have become dense along the real axis from 0 to $2\hbar\omega_L$. By analogy with similar integrals which have been calculated in other connections, we can sketch (Fig. 3) the real and imaginary parts of $F(E-i\delta)$, where δ is an infinitesimal positive number. The qualitative features of Fig. 3 which are certainly independent of the crystal model used are the general behavior of the real part for $E<0$ and $E>2\hbar\omega_L$; the fact that the imaginary part is positive for $0<E<2\hbar\omega_L$ and vanishes outside this range, and the rapid increase

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

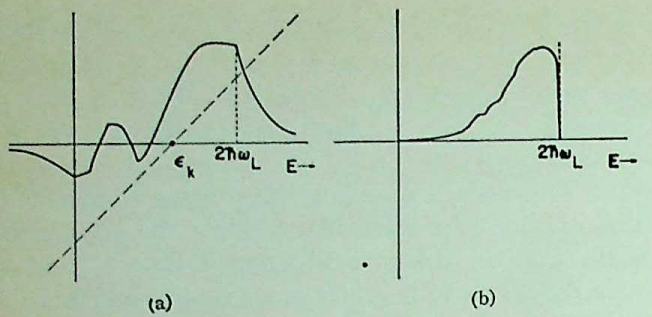


FIG. 3. Real (a) and imaginary (b) parts of $F(E-i\delta)$ for $N\to\infty$. The dashed line in (a) is $E-\epsilon_k$. Notice that, in this model, for a very large anharmonic term, the localized mode (defined now by the intersection) is an eigenstate, since then the imaginary part of F vanishes. For a realistically small anharmonic force, however, the intersection will not be far removed from ϵ_k.

in general of the absolute values of the real and imaginary parts of $F(E)$ as E increases from 0 to $2\hbar\omega_L$. This latter follows from the rapid increase of the density of states of two phonons which is a factor in $\rho(k,\epsilon)$.

$D(E)$ and hence the denominator of the integrand of Eq. (21) can never have any roots but real ones, because of the Hermiticity of the Hamiltonian. We can therefore contract the contour around the cut and obtain

$$(\Psi(0),\Psi(t))=\frac{1}{\pi}\int^{2\hbar\omega_L} e^{-iEt}\frac{h(E)dE}{[E-\epsilon_k-g(E)]^2+h^2(E)}, \quad (23)$$

where we have set

$$F(E-i\delta)=g(E)+ih(E). \quad (24)$$

The form of Eq. (23) is such that if $h(E)$ is small and $g(E)$ is slowly varying in the neighborhood of E_0 where E_0 is chosen such that

$$E_0-\epsilon_k-g(E_0)=0, \quad (25)$$

then it becomes approximately

$$(\Psi(0),\Psi(t))=e^{-iE_0t}e^{-\Gamma t/2} \quad (26)$$

with

$$\Gamma/2=h(E_0)=\pi\rho(k,E_0). \quad (27)$$

This is the same result which would be obtained from first-order perturbation theory, except there E_0 would be replaced by ϵ_k. To obtain a numerical answer from Eq. (27) one needs first to find out what B_{klm} is implied by a reasonable form of C_{ij} in Eq. (2); this is not a trivial task because it involves finding the eigenvectors of the secular matrix of the imperfect harmonic crystal.

V. CALCULATION OF $F(E)$

We shall now take a simple model for the harmonic crystal with a single impurity, for which we already know the eigenvectors,[10] and calculate $F(E)$ as expressed in Eq. (14).

[10] W. M. Visscher, Phys. Rev. 129, 28 (1963).

A simple and perhaps fairly reasonable form for the anharmonic energy is

$$H'=b(x_{000}-x_{100})^3-b(x_{000}-x_{-100})^3, \quad (28)$$

where x_{ijk} is the displacement in the x direction of the atom in the (ijk) simple cubic lattice site from its equilibrium position. In the model that we are using motions in the 3 Cartesian direction are independent and separable, so Eq. (28) represents a general central cubic anharmonic interaction between the impurity at (000) and its nearest neighbors. For positive b, H' gives rise to an extra repulsion for small interatomic distances, an extra attraction for large ones. It can be thought of as representing a physical situation where the impurity is a misfit in an otherwise perfect harmonic crystal.

If we now expand x_{ijk} in terms of the creation-annihilation operators $a_s^\dagger$, a_s according to

$$x_{ijk}=\sum_s (C_{s,ijk}{}^\dagger a_s{}^\dagger+C_{s,ijk}a_s)=\sum_s x_{ijk}{}^s, \quad (29)$$

(where the sum is over the normal modes of the imperfect harmonic crystal) and substitute into Eq. (28), we can pick out the terms having the form of those contained in h [Eq. (1)].

$$h=3b\sum_{l,m}[(C_{l,000}-C_{l,100})(C_{m,000}-C_{m,100})-(C_{l,000}-C_{l,-100})(C_{m,000}-C_{m,-100})]\times(C_{k,000}{}^\dagger-C_{k,100}{}^\dagger)a_l{}^\dagger a_m{}^\dagger a_k. \quad (30)$$

Here, as before, l and m refer to normal modes in the quasicontinuum of the imperfect harmonic crystal (phonons), and k refers to the localized mode.

In arriving at Eq. (30) we have made use of the fact that the localized mode is always an even excitation of the lattice about the impurity position, i.e., $C_{k,100}=C_{k,-100}$. If we were to use boundary conditions such that the phonons, too, are either even or odd with respect to the impurity position, then the contributions to the sum in Eq. (30) will have (l even, m odd) or (l odd, m even). Thus

$$B_{klm}=6b(C_{l,000}-C_{l,100})(C_{m,-100}-C_{m,100})\cdot(C_{k,000}{}^\dagger-C_{k,100}{}^\dagger)a_l{}^\dagger a_m{}^\dagger a_k \quad (31)$$

for l corresponding to an even mode; B_{klm} vanishes otherwise.

Specializing now to the case of a mass impurity (no change in the harmonic forces connecting the impurity to its neighbors) we can calculate the C's.

Equation (5′) of Ref. 10 may be written

$$x_{ijk}{}^s=g_{ijk,000}{}^s\delta\mu\omega_s{}^2x_{000}{}^s, \quad (32)$$

where $g^s=g(\omega_s{}^2)$ is given there in Eq. (23′), and $\delta\mu$ is the difference between the impurity and host mass. Equation (32) implies that the frequencies of the normal modes for which the displacement of the impurity atom

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

does not vanish are given by the roots of

$$g_{000,000^s}\delta\mu\omega_s^2=1\,, \tag{33}$$

and that the C's of Eq. (31) satisfy

$$C_{s,\alpha}=(g_{\alpha 0^s}/g_{00^s})C_{s,0} \tag{34}$$

for the even modes, and for the odd modes the C's are the same as for the perfect crystal. In Eq. (34) we have denoted the triplet (ijk) by α. For the even modes, then, a complete specification of the C's awaits only a normalization condition, which, upon requiring that

$$[p_\alpha,x_{\alpha'}]=-i\hbar\delta_{\alpha\alpha'} \tag{35}$$

and

$$[a_s,a_{s'}{}^\dagger]=\delta_{s's}\,, \tag{36}$$

can be straightforwardly shown to be

$$|C_{s,0}|^2=\frac{\hbar}{2\omega_s\mu}(g_{00^s})^2\Big/\left[\sum_\gamma(g_{\gamma 0^s})^2+\frac{\delta\mu}{\mu}(g_{00^s})^2\right]. \tag{37}$$

Then, if we use the fact that

$$g_{00^s}-g_{10^s}=2(\omega_s/\omega_L)^2g_{00^s}+(2/\mu\omega_L{}^2)\,, \tag{38}$$

which can be verified from the forms of the g's given in Eq. (25′) of Ref. 10, Eq. (31) can be rewritten

$$|B_{klm}|^2=(24b)^2(\mu'/\mu)^4k^4l^4|C_{k,0}|^2|C_{l,0}|^2 \times|C_{m,1}-C_{m,-1}|^2, \tag{39}$$

where $\mu'=\mu+\delta\mu$ and $l=\omega_l/\omega_L$, $k=\omega_k/\omega_L$, and where m refers to an odd mode in which the impurity is located at a node. It then follows that $C_{m,\alpha}$ is not different from that for the perfect lattice, and therefore

$$|C_{m,1}-C_{m,-1}|^2=(\hbar/2\mu N\omega_m)4\sin^2\varphi_x{}^m, \tag{40}$$

where N is the number of atoms in the crystal and $\varphi_x{}^m$ is the x component of the wave vector of the mth phonon.

We are now in a position to write down an integral form for the sum in $F(E)$, Eq. (14). The sum over l can be replaced by an integral as follows:

$$\sum_l|C_{l,0}|^2\rightarrow\frac{1}{2\pi i}\int_C\frac{\hbar}{2\omega_l}\frac{d\omega_l{}^2}{(g_{00}{}^l)^{-1}-\delta\mu\omega_l{}^2}\rightarrow \frac{\hbar}{2\pi\gamma}\int_0^{\omega_L}d\omega_lG(l)\,, \tag{41}$$

where the contour encircles the zeroes of the denominator in the range $0<\omega_l<\omega_L$, and the second step is achieved upon taking the limit $N\rightarrow\infty$ and contracting the contour about the cut which then appears from 0 to ω_L.

$$G(l)=2\gamma\frac{\mathrm{Im}g_{00}{}^l(\omega_l{}^2-i\epsilon)}{(1-\delta\mu\omega_l{}^2\,\mathrm{Re}g_{00}{}^l)^2+(\delta\mu\omega_l{}^2\,\mathrm{Im}g_{00}{}^l)^2}\,, \tag{42}$$

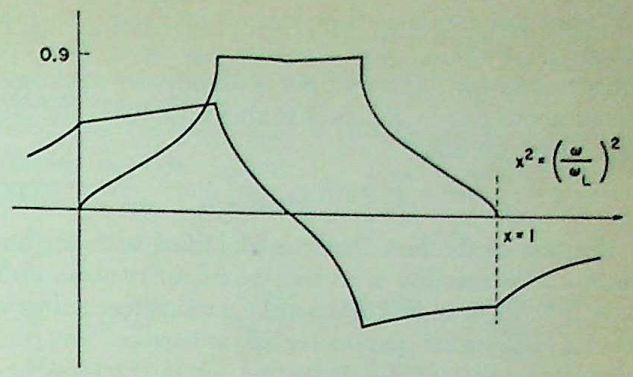


FIG. 4. Real and imaginary parts of $2\gamma g_{00}(\omega)$.

and γ is the harmonic force constant. The sum over m can be similarly expressed:

$$\sum_m\frac{\hbar}{2\mu N\omega_m}\rightarrow\frac{\hbar}{2\pi\gamma}\int_0^{\omega_L}d\omega_mG_0(m)\,, \tag{43}$$

where

$$G_0(m)=2\gamma\,\mathrm{Im}g_{00}(\omega_m{}^2-i\epsilon)\,. \tag{44}$$

Now, if we take E to be removed from the real axis, we can write

$$F(E)=2\left(\frac{24b\hbar\omega_L}{\pi\gamma}\right)^2\left(\frac{\mu'}{\mu}\right)^4k^4|C_{k,0}|^2 \cdot\int_0^1\frac{dldml^4G(l)G_0(m)\langle\sin^2\varphi_x{}^m\rangle}{E-\epsilon_l-\epsilon_m}\,, \tag{45}$$

where the bracket denotes an average over all modes with energy ϵ_m. The dimensionless function $G(l)$ has been calculated numerically in Ref. 10; its shape for various impurity masses is exhibited in Fig. 9 of that paper and Fig. 4 of the present paper shows the real and imaginary parts of g_{00}. $|C_{k,0}|^2$ is, as can be seen from Eq. (29), just the contribution of the localized mode to the mean-square displacement of the impurity; this has also been calculated in Ref. 10. We can, semiquantitatively, identify

$$bC_{k,0}/\gamma=r \tag{46}$$

as the ratio of the anharmonic potential energy of the impurity atom to its harmonic energy when the localized mode is excited.

The imaginary part of $F(E)$, according to Eq. (27), is equal to the half-width of the localized mode line.

$$\Gamma/2=h(E)=\frac{2\hbar\omega_L}{\pi}(24r)^2\left(\frac{\mu'}{\mu}\right)^4k^4 \cdot\int_0^1dll^4G(l)G_0\left(\frac{E}{\hbar\omega_L}-l\right)\left\langle\sin^2\varphi_x\left(\frac{E}{\hbar\omega_L}-l\right)\right\rangle. \tag{47}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

An approximate numerical integration[11] for the case where $E=1.12\hbar\omega_L$, which is the harmonic localized mode energy for this model if $\mu'=0.5\,\mu$, yields the result that the integral in Eq. (47) is about 5×10^{-3}. For this case, then, we find

$$\Gamma/2\approx0.2\hbar\omega_L r^2. \tag{48}$$

Because of the fact that r is identified with the anharmonic interaction of an isolated impurity atom with its neighbors, it is difficult to make a numerical estimate for it. The usual parameter of anharmonicity, the Grüneisen constant,[12] is measured for interactions between atoms of the host crystal. But if we associate these quantities with each other, we find that a reasonable Grüneisen constant (of the order of unity) for a crystal implies that r for the force between each atom and its nearest neighbors is of the order of 10^{-2}. Therefore, from Eq. (48) we see that the order of magnitude of the ratio of the half-width to energy of the anharmonic localized mode is 10^{-4}. Within the limitation of the calculations, this result is not in disagreement with the widths obtained in the references cited above.[6] No careful quantitative work will be done here, because we do not believe the model we use to be realistic enough to warrant it. The significance of the present work lies more in the form than in the magnitude of the answers.

The real part of $F(E)$ is just Eq. (45) with the integral replaced by a principal value integral. It will be of the same order-of-magnitude as the imaginary part; hence the shift of the localized mode (downward in energy for frequencies close to the phonon cutoff) is comparable to its width.

[11] An analytic approximation can be made in the limit of small impurity mass.

[12] See, for example, P. G. Klemens, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1958), Vol. 7, p. 1.

VI. ACKNOWLEDGMENT

We are indebted to Dr. A. A. Maradudin for stimulating and pertinent correspondence and discussions.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Energy Bands for the Iron Transition Series*

L. F. Mattheiss†
Solid-State and Molecular Theory Group, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 16 December 1963)

Preliminary energy-band calculations for elements of the iron transition series have been completed using the augmented plane-wave method. The results include plots of energy as a function of wave vector along a line of symmetry for elements crystallizing in the face-centered cubic (Ar, Co, Ni, Cu), body-centered cubic (V, Cr, Fe), and hexagonal close packed (Ti, Zn) structures. These results indicate the presence of systematic trends in the band structures for the various elements and provide some justification for the application of the rigid band model to transition metals and their alloys.

I. INTRODUCTION

ENERGY bands have been calculated for a majority of the elements in the iron transition series using the augmented plane-wave (APW) method.[1,2] While the present results are preliminary in nature and not in any sense complete, they may be of some interest to experimentalists and theoreticians who are concerned with the electronic structure of the transition-series elements. The present results represent energy bands for three different crystal structures, with a variety of lattice constants. Despite the detailed differences that are imposed by symmetry requirements and variations in lattice constants, the results suggest some interesting and rather clear-cut trends in the band structure of these elements as one proceeds through the transition series. These calculations lend some support to the rigid band model for the transition series. They support the hope that systematic studies of the band structure of the transition-series elements can provide useful qualitative, and perhaps quantitative, information concerning their electronic structure.

As in all calculations involving d electrons, the results are sensitive to the choice of potentials. The crystal potentials used in these calculations were all constructed in an analogous manner, and were approximated by a superposition of atomic potentials. The method involves the use of Hartree-Fock solutions to the corresponding atomic problem[3] and the free-electron-exchange approximation.[4] The details of this method for constructing approximate crystal potentials have been described earlier,[5] though a brief resumé is presented in Sec. II of

* This work was supported by the National Science Foundation.

† Presently employed at Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey.

[1] J. C. Slater, Phys. Rev. 51, 846 (1937).

[2] J. H. Wood, Phys. Rev. 126, 517 (1962).

[3] R. E. Watson, Phys. Rev. 119, 1934 (1960); R. E. Watson and A. J. Freeman, Phys. Rev. 123, 521 (1961).

[4] J. C. Slater, Phys. Rev. 81, 385 (1951).

[5] L. F. Mattheiss, Phys. Rev. 133, A1399 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

this paper, along with other information pertaining to this present series of calculations. The energy bands along a single line of symmetry in the appropriate Brillouin zone are presented in Sec. III for Ar, Ti, V, Cr, Fe, Co, Ni, Cu, and Zn, while the last section contains a brief discussion of these results.

II. DESCRIPTION OF THE CALCULATIONS

In these calculations, the crystal potential has been approximated by a superposition of atomic potentials. The Coulomb and exchange contributions to the crystal potential are treated separately. An approximate crystal Coulomb potential and charge density in a given atomic cell is obtained by expanding the neutral atom Coulomb potentials and charge densities of neighboring atoms about the origin, using Löwdin's alpha function expansion,[6] keeping only the $l=0$ or spherically symmetric terms in these expansions. Using the free-electron-exchange approximation, the exchange potenial is proportional to the cube root of the superimposed atomic-charge densities.

The potentials obtained by this method are generally rather flat near the boundaries of the atomic cell, at least in the case of metals, so they are readily approximated by a "muffin-tin" type potential, as required by the APW method. The constant value of the potential outside the APW spheres is taken as the average value of the potential in this region. This usually results in a discontinuity in the potential at the sphere radius amounting to a few hundredths of a Rydberg.

In the construction of approximate crystal potentials for transition-series elements, there is frequently some ambiguity in choosing the most reasonable atomic configuration. This sort of difficulty can only be answered satisfactorily by experimental information and/or self-consistent energy-band calculations. For the present, we have been content to study the effect that changing the atomic configuration has on the band structure. In addition, there are magnetic effects which create additional complications in this series of elements. For simplicity, all magnetic effects have been neglected in these calculations, and the crystals have been assumed to be nonmagnetic in character.

The lattice constants which have been used in these calculations have generally been the room-temperature values as tabulated by Pearson.[7] The exceptions are those for Ar and Zn. The lattice constant for Ar is the low-temperature value obtained by Dobbs and Jones.[8] In the case of Zn, Harrison[9] has extrapolated the room-temperature lattice constants to low temperatures since the results are expected to be sensitive to the choice of c/a ratio. For purposes of comparison, his values have been used in these calculations. Table I contains a summary of the elements considered in these calculations, their structures, the values of the lattice constants, and the assumed atomic configurations.

TABLE I. In this table, we list the elements of the iron transition series, their structures, the lattice constants used in this series of calculations (in atomic units), and the assumed atomic configurations.

Element	Structure	a(au)	c(au)	Configuration
Ar	fcc	10.0346		$(3s)^2(3p)^6$
K				
Ca				
Sc				
Ti	hcp	5.5755	8.8503	$(3d)^3(4s)^1$
V	bcc	5.7225		$(3d)^4(4s)^1$
Cr	bcc	5.4512		$(3d)^5(4s)^1$
Mn				
Fe	bcc	5.4168		$(3d)^7(4s)^1$
Co	fcc	6.6975		$(3d)^8(4s)^1$
Ni	fcc	6.6590		$(3d)^9(4s)^1$
Cu	fcc	6.8309		$(3d)^{10}(4s)^1$
Zn	hcp	5.0120	9.1453	$(3d)^{10}(4s)^2$

III. RESULTS

The principal results of these calculations are presented in Fig. 1. These results represent plots of energy as a function of wave vector along lines of symmetry from the center to a boundary of the appropriate Brillouin zones. For the face-centered cubic structure (Ar, Co, Ni, and Cu), the bands are plotted from Γ along the Δ direction to the point X, using the notation of Bouckaert, Smoluchowski, and Wigner.[10] In the body-centered cubic structure (V, Cr, and Fe), the bands are plotted from Γ along the Δ direction to the point H. Finally, in the hexagonal close-packed structure (Ti and Zn), they start at Γ and proceed along the line T in the $k_z=0$ plane which terminates at the point K, one of the vertices of the hexagon (in the notation of Herring[11]).

The energy is in Rydbergs and the wave vectors for the different elements are drawn to scale for purposes of comparison. The horizontal dashed lines represent rough estimates of the Fermi energy for each element. For simplicity, some of the more highly excited bands have been omitted in some cases, particularly in the face-centered cubic structure, or in other situations, they have been sketched in by dashed lines.

In Fig. 2, it is shown what effect varying the atomic configuration has on the band structure of a typical element, namely vanadium. The bands to the right are the ones shown in Fig. 1 for vanadium; the ones to the left are those obtained from a potential which results from an atomic configuration containing an additional $4s$ electron and one less $3d$ electron. In general, this results in a narrowing of the $3d$ band and a decrease in

[6] P. O. Löwdin, Advan. Phys. 5, 1 (1956).
[7] W. B. Pearson, *A Handbook of Lattice Spacings and Structures of Metals and Alloys* (Pergamon Press, Inc., New York, 1958).
[8] E. R. Dobbs and G. O. Jones, in *Reports on Progress in Physics*, edited by A. C. Stickland (The Physical Society, London, 1957), Vol. 20, p. 516.
[9] W. A. Harrison, Phys. Rev. 126, 497 (1962).
[10] L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).
[11] C. Herring, J. Franklin Inst. 233, 525 (1942).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

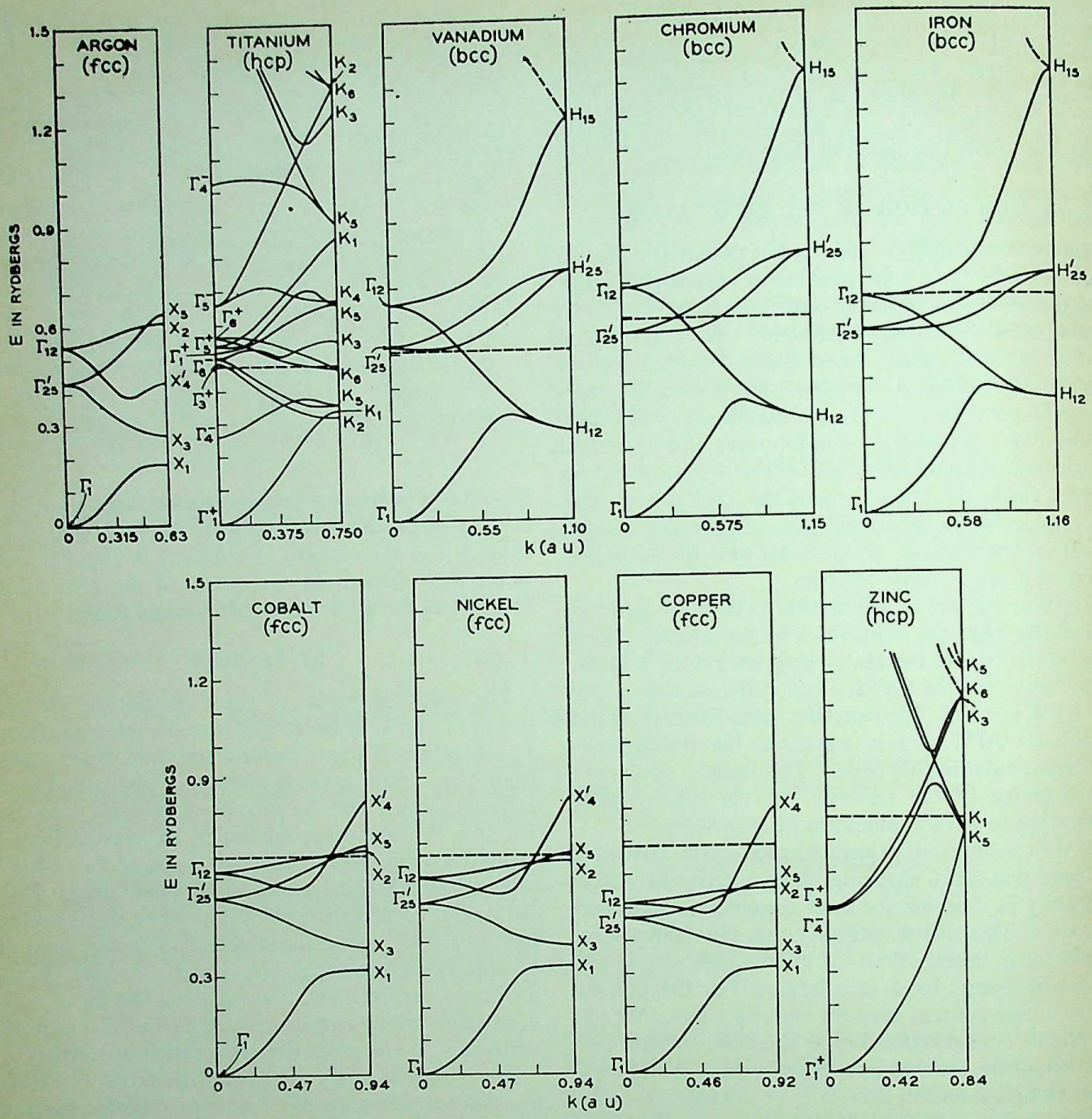


FIG. 1. Energy bands for Ar, Ti, V, Cr, Fe, Co, Ni, Cu, and Zn as a function of wave vector along a line of symmetry in the appropriate Brillouin zone. For the face-centered cubic structure, the bands are plotted from Γ along Δ to the point X. For the body-centered cubic structure, they are plotted from Γ along Δ to the point H. Finally, for the hexagonal close-packed structure, they are plotted from Γ along T to the point K. The energies are in Rydbergs and the wave vectors are in atomic units.

the energy separation between the top of the $3d$ band and the bottom of the $4s$-$4p$ bands.

IV. DISCUSSION

In the simplified picture of the energy bands for the iron transition-series elements, one finds a narrow $3d$ band in the midst of a rather broad $4s$-$4p$ band. The width of the $3d$ band and especially its position relative to the bottom of the $4s$-$4p$ band depend rather critically on the potential. Nevertheless, the results of Fig. 1 exhibit a reasonably smooth variation from element to element, especially for those substances having the same crystal structure. This seems to lend some support to the rigid band model for the transition-series elements, an approximation which has been of considerable value in understanding the electronic properties of these elements and their alloys.

There is a gradual narrowing of the $3d$ band as one progresses through the series. This effect was discussed by Slater in order to explain the occurrence of ferromagnetism in the latter part of the series.[12] In going from Cu to Zn, the $3d$ band suddenly drops about 0.5 Ry below the bottom of the $4s$-$4p$ bands, and its width

[12] J. C. Slater, Rev. Mod. Phys. 25, 199 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

decreases to less than 0.1 Ry. As a result, the energy bands for zinc are very free-electron-like. For those elements where the 3*d* band falls in the middle of the 4*s*-4*p* bands, the interactions between states having the same symmetry causes considerable modification to the free-electron bands, though at points of symmetry, the effect is sometimes small. The bands for Ti and Zn demonstrate this effect nicely.

The results of Fig. 2 emphasize the uncertainty which is inherent in any energy-band calculation for a transition-series element. These uncertainties have been pointed out previously in the literature, particularly by Callaway.[13] These difficulties can only be cleared up satisfactorily with the aid of more detailed experimental information regarding the band structure of these elements in addition to self-consistent energy-band calculations.

The results presented here are not complete enough to permit detailed comparisons to be made with experiment or a discussion of the resulting Fermi surfaces. However, there are some striking similarities between the energy bands shown in Fig. 1 and the results obtained by earlier calculations. In particular, there is good qualitative agreement between the Cu results shown in Fig. 1 and the bands calculated by Segall[14] and also by Burdick.[15] Similarly, the results for Fe are in good agreement with the published results of Wood.[2] The agreement for Ar with the results of calculations by Knox and Bassani[16] is good, and has been described previously.[5] In the case of Cr, it is difficult to compare the present results with those of earlier calculations by Asdente and Friedel[17] since they neglect the interactions between the 3*d* band and the 4*s*-4*p* bands. However, the present results do justify, to some extent, the treatment of Cr by Lomer,[18] who used the results of Wood's iron calculations to discuss the energy bands in antiferromagnetic Cr.

In the case of Zn, the ordering of levels is identical with that obtained by Harrison.[9] This ordering differs from that obtained in earlier calculations for hexagonal close-packed metals by Herring and Hill for Be[19] and Falicov for Mg.[20] This change in ordering might be due to the presence of an occupied 3*d* band just below the 4*s*-4*p* bands. The results for Ti agree qualitatively with those obtained by Altmann and Bradley.[21] Finally, the Ni results are in good agreement with those obtained by Hanus.[22]

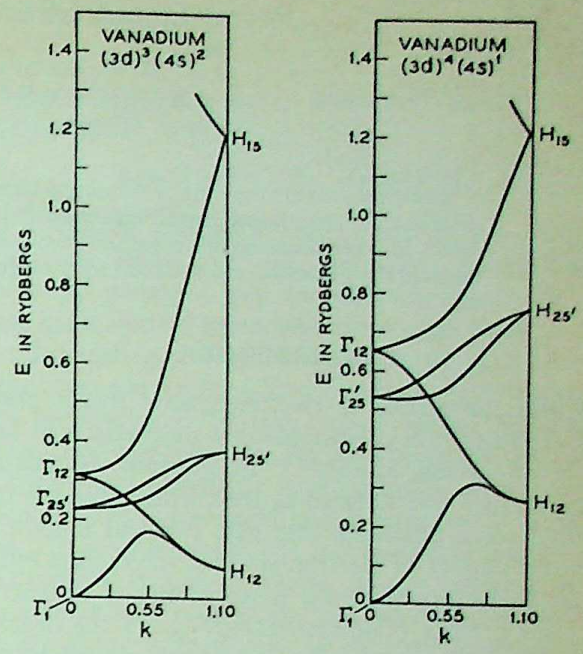


FIG. 2. Energy bands for vanadium along the line Δ obtained from two different potentials. The bands to the left resulted from an atomic configuration of $(3d)^3(4s)^2$ while those to the right involved a $(3d)^4(4s)^1$ configuration.

ACKNOWLEDGMENTS

The author wishes to thank Professor John C. Slater for encouraging and supporting this work. The author is grateful to J. H. Wood and A. C. Switendick for the use of their APW programs and for many useful discussions and comments.

[13] J. Callaway, Phys. Rev. **99**, 500 (1955).
[14] B. Segall, Phys. Rev. **125**, 109 (1962).
[15] G. A. Burdick, Phys. Rev. **129**, 138 (1963).
[16] R. S. Knox and F. Bassani, Phys. Rev. **124**, 652 (1961).
[17] M. Asdente and J. Friedel, Phys. Rev. **124**, 384 (1961).
[18] W. M. Lomer, Proc. Phys. Soc. (London) **80**, 489 (1962).
[19] C. Herring and A. G. Hill, Phys. Rev. **58**, 132 (1940).
[20] L. M. Falicov, Phil. Trans. Roy Soc. **A255**, 55 (1962).
[21] C. J. Bradley, Ph.D. thesis, Oxford University, 1962 (unpublished).
[22] J. G. Hanus, Solid-State and Molecular Theory Group, Massachusetts Institute of Technology, Quarterly Progress Report No. 44, p. 29, 1962 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Fermi-Liquid Effects in Cyclotron Resonance

P. M. Platzman and K. C. Jacobs
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received 19 November 1963)

The steady-state properties of transverse circularly polarized waves propagating along a static magnetic field in a uniform plasma (metal) are considered. Using the Landau theory of Fermi liquids, we compute the effect of correlations on the reflection properties of a semi-infinite metallic plasma. The expressions are numerically evaluated and discussed for a range of parameters pertinent to the alkalis.

I. INTRODUCTION

THE properties of electromagnetic waves propagating in a plasma along a magnetic field have been extensively studied.[1–3] In previous calculations the electrons were treated as the current carrying constituent of a noninteracting gas. For real metals the *a priori* neglect of electron-electron interaction effects is unsatisfactory since the mean energy of interaction of the electrons (Coulomb energy) is of the order of their mean kinetic energy. Under such conditions substantial correlation effects in the motion of the electrons may be expected. From a practical point of view, however, it is known that for a wide range of physical parameters, treatments which neglect explicit correlation effects provide an accurate description of the electromagnetic properties of metals and semimetals. We are interested in determining the range of physical parameters, frequency, external magnetic field, carrier density, etc., where the effect of electron-electron correlation may, hopefully, be measured experimentally.

In 1956, Landau[4] constructed a semiphenomenological theory of a system of fermions, such as He_3, interacting via a short-range two-body force. This work was later extended by Silin[5] to a Fermi liquid with long-range Coulomb interactions. We make use of the Fermi-liquid theory to compute the reflection properties of a semi-infinite slab of metal placed in a magnetic field oriented perpendicular to its surface. In this geometry for electromagnetic waves incident normally, it is known[1,3] that the metallic sample will exhibit an absorption edge at approximately the Doppler shifted cyclotron frequency.

$$\omega_c/\omega = 1+(q/q_0)(V_F/c)\,, \tag{1}$$

where ω_c is the electron cyclotron frequency, q is, crudely speaking, the wave number of the electromagnetic field in the medium, q_0 the wave number in free space, and V_F is the Fermi velocity. The resonance in the electrons motion which exhibits itself as a rapid fluctuation in the surface impedance of the sample is shifted from the value $\omega_c/\omega=1$. This comes about since the electrons traveling with the phase velocity of the wave see a Doppler shifted frequency. It is these electrons which interact most strongly with the wave. As a result the resonance in the impedance is shifted. Specifically we are interested in possible shifts in the position of, and modifications of the shape of the Doppler shifted electron-cyclotron resonance as the result of electron–electron interactions.

II. CALCULATION

The basic assumption of the Fermi-liquid theory is that it presupposes that as the interaction is turned on, the single-particle states in the neighborhood of the last occupied one remain approximate eigenstates of the interacting system, and that there is a one-to-one correspondence between these states and the single-electron states of the noninteracting Fermi gas. These approximate eigenstates are called quasiparticle states. For slowly varying external disturbances, the transport properties of the Fermi liquid are completely described by the quasiparticle distribution function $f(\mathbf{P},\mathbf{X})$ in momentum and configuration space.[6] (The external disturbance varies slowly enough in space so that the lack of commutivity of $\mathbf{P}$ and $\mathbf{X}$ is unimportant.) In equilibrium and at zero temperature the distribution of quasiparticles $f_0(\mathbf{P},\mathbf{X})$ is

$$\begin{aligned} f_0(\mathbf{P},\mathbf{X}) &= 1\,, \qquad |P|<P_F\,, \\ &= 0\,, \qquad |P|>P_F\,. \end{aligned} \tag{2}$$

For the spherically symmetric electron gas the quasiparticle energy is $E^0(P)=P^2/(2m^*)$, where m^* is an effective mass whose value depends on the dynamics of the interaction between quasiparticles. In nonequilibrium situations, the quasiparticle distribution function satisfies a transport equation similar to the Boltzmann equation. We are interested in the transport equation in the linear approximation, i.e., when the deviation from f_0 is small. If we write $f=f_0+f_1$, with

[1] P. B. Miller and R. R. Haering, Phys. Rev. **128**, 126 (1962).
[2] P. M. Platzman and S. J. Buchsbaum, Phys. Rev. **128**, 1004 (1962).
[3] P. M. Platzman and S. J. Buchsbaum, Phys. Rev. **132**, 2 (1963).
[4] L. D. Landau, Zh. Eksperim. i Teor. Fiz. **30**, 1058 (1956) [English transl.: Soviet Phys.—JETP **3**, 920 (1956)].
[5] V. P. Silin, Zh. Eksperim. i Teor. Fiz. **33**, 495 (1957) [English transl.: Soviet Phys.—JETP **6**, 387 (1958)].

[6] The distribution function $f(\mathbf{P},\mathbf{X})$ is a matrix in the spin variables $F(\mathbf{P},\mathbf{X})=f(\mathbf{P},\mathbf{X})+\mathbf{m}(\mathbf{P},\mathbf{X})\cdot\boldsymbol{\sigma}$. The function $\mathbf{m}(P,x)$ is pertinent only when quantities which depend on the spin (the susceptibility) are of interest. We will be concerned only with $f(\mathbf{P},\mathbf{X})$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$f_1 \equiv \delta[E^0(P)-E_F]g$ the linearized transport equation for g becomes

$$\partial g/\partial t+\mathbf{V}\cdot\nabla_X(g+E_1)+(e/c)(\mathbf{V}\times\mathbf{H})\cdot\nabla_P(g+E_1) -e\mathbf{E}\cdot\mathbf{V}=-\nu_c g. \quad (3)$$

The quantity ν_c is a phenomenological short-range collision term which permits the system to relax to equilibrium; $V_\alpha \equiv \partial E^0(P)/\partial P_\alpha = P_\alpha/m^*$ is the velocity of the quasiparticles and

$$E_1=\int S(\mathbf{P},\mathbf{P}')\delta[E_F-E^0(P')]g(\mathbf{P}',\mathbf{X}) \times[d^3P'/(2\pi\hbar)^3] \quad (4)$$

is the change in energy of the quasiparticles produced by a change in their distribution function.

The correlation, or interaction function $S(\mathbf{P},\mathbf{P}')$ is the basic quantity characterizing the Fermi liquid. Owing to the presence of the delta function in the distribution function f_0 and in the transport equation, the value of $S(\mathbf{P},\mathbf{P}')$ is only of interest for $|P|=|P'|=P_F$, where P_F is the Fermi momentum. $S(\mathbf{P},\mathbf{P}')$ may then be considered to be a function of the angle between $\mathbf{P}$ and $\mathbf{P}'$. It is conveniently represented by an infinite sequence of Legendre polynomials.

$$S(\hat{P}\cdot\hat{P}')=\textstyle\sum_n S_n P_n^{(0)}(\cos\theta)\,, \quad (5)$$

where $\hat{P}$ is a unit vector in the direction of $\mathbf{P}$. For convenience, we define a set of dimenionsless quantities

$$\delta_0 \equiv S_0 m^* P_F/(2\pi)^2\hbar^3$$
$$\delta_n \equiv [S_n m^* P_F/(2\pi)^2\hbar^3][(n-1)!/(n+1)!]n\geq 1. \quad (6)$$

The correlation function $S(\hat{P}\cdot\hat{P}')$ contains all the information pertaining to explicit dynamical many-body effects. The limit $S\rightarrow 0$ is the free electron or Hartree limit. In this limit the transport equation, Eq. (3) becomes the usual Vlasov equation and the only "correlation" effect which is included is the self-consistent electromagnetic field.

The correlation function is as important to the many electron problem as the shape of the Fermi surface is in the one-electron problem. Unfortunately, the function $S(\hat{P}\cdot\hat{P}')$ (for the region of metallic densities), cannot be computed from first principles.[7] Hopefully then a transport experiment of the kind to be considered here might enable one to measure the characteristics of the scattering function $S(\hat{P}\cdot\hat{P}')$. Ideally one would like to perform an experiment in which a specific moment of the scattering function is measured. It was first pointed out by Landau[4] that the zeroth and first moments of the scattering function were simply related to two elementary properties of the quasiparticle gas. The clothed mass m^* of the quasiparticles is related to the first moment S_1 by

$$m^*/m=1+\tfrac{1}{3}\delta_1 \quad (7)$$

and the velocity of ordinary sound v_S is related to the zeroth moment S_0 by

$$v_S=(\delta_0+1)^{1/2}[P_F/(3mm^*)^{1/2}]. \quad (8)$$

A cyclotron resonance experiment in the extreme anomalous limit, in the so-called Azbel'-Kaner geometry,[8] measures only the effective mass.[9] On the other hand, the cyclotron resonance experiment in the so-called Galt geometry[10] involves, as we shall see, all of the moments. In the Galt geometry a static magnetic field perpendicular to the surface of the sample is utilized and the reflection (or absorption) of circularly polarized electromagnetic waves propagating along the magnetic field is measured as a function of the applied field. The solution, neglecting correlations, of the coupled Maxwell-Boltzmann equations in the presence of a single boundary (assuming specular reflection of the carriers at the surface) may be reduced to the solution of an equivalent infinite medium problem.[3] The final expression for the surface impedance Z is

$$\frac{Z}{Z_0}=(2i/\pi)\int_0^\infty \frac{dq}{[q^2-\epsilon(q,\omega)]}, \quad (9)$$

where Z_0 is the impedance of free space. The quantity $\epsilon(q,\omega)$ is the finite wave number and frequency-dependent dielectric constant for the infinite medium, $[\epsilon(q,\omega)\equiv 1+\sigma^+(q,\omega)/i\omega]$.

The quantity E_1 is Eq. (4) is related to a first-order change in the distribution function by the relationship

$$E_1=\int S(\mathbf{P},\mathbf{P}',\mathbf{X},\mathbf{X}')\delta[E_F-E^0(P')]g(\mathbf{P}',\mathbf{X}') \times\frac{d^3P'd^3X'}{(2\pi)^3\hbar^3}. \quad (10)$$

In writing Eq. (4) we have assumed that $S(\mathbf{P},\mathbf{P}',\mathbf{X},\mathbf{X}')=S(\mathbf{P},\mathbf{P}')\delta(\mathbf{X}-\mathbf{X}')$. With this assumption the spatial part of the problem is identical with that in the noninteracting case. For an interacting gas of quasiparticles in real materials correlations are not local in space but do in fact extend over distances of the order of a Debye length ($\lambda_D \simeq V_F/\omega_P$). For metallic densities this is of the order of the interparticle spacing. The Debye length is small compared with distances over which the field varies appreciably, typically distances of the order of the high-frequency classical skin depth ($\lambda_S \simeq c/\omega_P$), so that to a good approximation we can neglect nonlocal effects. With the locality assumption on the function

[7] C. Herring has recently computed the scattering function S as a power series in r_s ($r_s \equiv n^{1/3}\,a_H$ where n is the electron density and a_H is the Bohr radius) for metals $r_s \approx 2$–6 (private communication).

[8] M. Ya. Azbel, E. A. Kaner, Zh. Eksperim. i Teor. Fiz. 30, 811 (1956) [English transl.: Soviet Phys.—JETP 3, 772 (1956)].

[9] This statement has been proved by J. M. Luttinger (private communication).

[10] J. K. Galt, W. A. Yager, F. R. Merritt, B. B. Cetlin, and A. D. Brailsford, Phys. Rev. 114, 1396 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

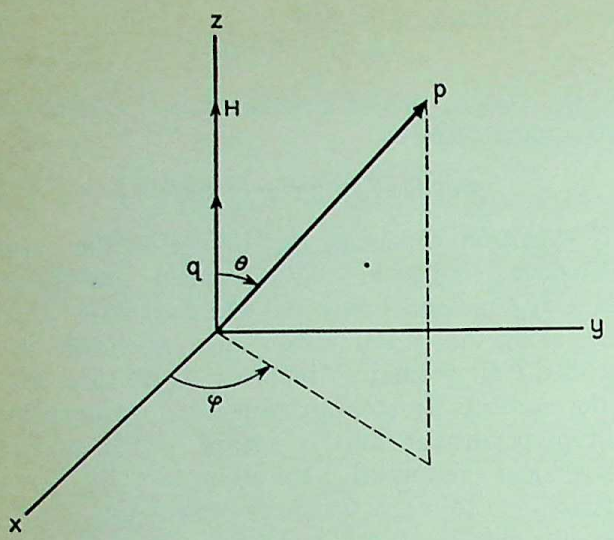


FIG. 1. Polar coordinate system used for solution of the transport equation, Eq. (3). H is the magnetic field, q is the propagation vector, and P is the quasiparticle momentum.

$S(\mathbf{P},\mathbf{P}',\mathbf{X},\mathbf{X}')$, the expression for the impedance within the framework of the Fermi-liquid theory is still given by Eq. (9). However, the dielectric constant $\epsilon(q,\omega)$ will now include the effects of correlations.

This dielectric constant is obtained by linearizing the transport equation, Eq. (2), with an rf field varying as $\exp(i\mathbf{q}\cdot\mathbf{X}-\omega t)$ and solving for the induced current. The current is given by

$$I^\alpha(q)=\frac{ne}{m}\int\frac{d^3P}{(2\pi)^3\hbar^3}P^\alpha g(\mathbf{P},q)\delta(E_P{}^0-E_F) \qquad (11)$$

or, equivalently, using Eq. (4)

$$I^\alpha(q)=\frac{ne}{m^*}\int\frac{d^3P}{(2\pi)^3\hbar^3}P^\alpha(g+E_1)\delta(E_P{}^0-E_F)\,. \qquad (12)$$

We choose our coordinate system (see Fig. 1) so that the static magnetic field H and the propagation vector $\mathbf{q}$ point along the z axis. The angles θ and Φ are polar angles specifying the direction of $\mathbf{P}$. For a right-handed circularly polarized wave ($\mathbf{E}=E_x\hat{U}x+iE_y\hat{U}_y$) the transport equation for $g(\mathbf{P},q)$ is,

$$-i\tilde{\omega}g(\mathbf{P},q)+\frac{iqP}{m^*}\cos\theta g(P,q)+\omega_c{}^*\frac{\partial g}{\partial\Phi}$$
$$+\left[\frac{iqP}{m^*}\cos\theta+\omega_c{}^*\frac{\partial}{\partial\Phi}\right]$$
$$\times\int d^3P'/(2\pi\hbar)^3S(\mathbf{P},\mathbf{P}')\delta(E_F-E_P{}^0)g(\mathbf{P}',q)$$
$$\times eEe^{i\Phi}\frac{P}{m^*}\sin\theta, \qquad (13)$$

where $\omega_c{}^*=|eH/m^*c|$ and $\tilde{\omega}=\omega+i\nu_c$.

If we set $g(\mathbf{P},q)\equiv g(\theta,\Phi)=g(\theta)e^{i\Phi}$, then it is easily shown that

$$g(\theta)=\frac{ieEV_F{}^*\sin\theta}{\xi}+\frac{(qV_F{}^*\cos\theta+\omega_c{}^*)}{\xi}\Lambda(\theta)\,, \qquad (14)$$

where $V_F{}^*=P_F/m^*$ and

$$\xi=(\tilde{\omega}-\omega_c{}^*-qV_F{}^*\cos\theta)\,. \qquad (15)$$

The quantity $\Lambda(\theta)$ is

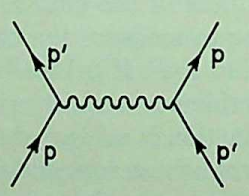
$$\Lambda(\theta)=\sum_{n=0}^{\infty}\delta_nP_n{}^{(1)}(\cos\theta)$$
$$\times\int_\pi^0 g(\theta')P_n{}^{(1)}(\cos\theta')d(\cos\theta')\,. \qquad (16)$$

Equation (14) is an infinite set of coupled linear equations for the quantities $\Omega_n=\int_\pi^0 P_n{}^{(1)}(\cos\theta)g(\theta)d(\cos\theta)$. If we terminate the spherical harmonic expansion for $S(\hat{P}\cdot\hat{P}')$ after a finite number of terms, then Eq. (16) may be solved self-consistently for the quantity $g(\theta)$. Knowing $g(\theta)$, we may substitute into Eq. (11) and evaluate the current.

For the purposes of this paper we will include the first three moments of $S(\hat{P}\cdot\hat{P}')$ S_0, S_1, and S_2. We have no real justification for terminating this series after so few terms; however, there are at least some crude arguments which suggest that a Legendre expansion for $S(\hat{P}\cdot\hat{P}')$ converges rapidly. Landau has shown that the function $S(\hat{P}\cdot\hat{P}')$ is proportional to the negative of the forward scattering amplitude of two quasiparticles. A typical type of exchange scattering diagram which enters into a microscopic calculation of $S(\hat{P}\cdot\hat{P}')$ is shown in Fig. 2. The dashed line is a bare Coulomb line and the solid lines represent electron (hole) propagators. This type of scattering diagram contributes a term to $S(\hat{P}\cdot\hat{P}')$ of the form

$$S_{\text{exch.}}(\hat{P}\cdot\hat{P}')\sim\frac{-e^2}{|\mathbf{P}'-\mathbf{P}|^2\epsilon_L(|\mathbf{P}'-\mathbf{P}|,0)}\,, \qquad (17)$$

where $\epsilon_L(q,0)$ is the zero-frequency longitudinal dielectric constant. If ϵ_L is computed by summing the



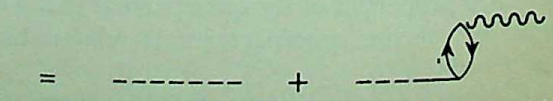


FIG. 2. A typical type of exchange scattering diagram which contributes to $S(\hat{P}\cdot\hat{P}')$. Solid lines are electron propagators and dashed lines represent Coulomb propagators.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

simple set of bubble diagrams shown in Fig. 2 (RPA approximation), we find

$$\epsilon_L(q,0)=1+(q_D^2/q^2)F(q)\,, \tag{18}$$

where $F(q)\approx 1$ for $q\ll q_F$ and $F(q)\approx\frac{1}{2}$ for $q=2q_F$. Using this expression for the dielectric constant, we find

$$S(0)/S(\pi)=\tfrac{1}{2}[1+8P_F^2/q_D^2]\,. \tag{19}$$

For potassium $S(0)/S(\pi)\approx 0.7$ so that on the basis of this crude computation the scattering is roughly spherical.

If we truncate Eq. (5) after three terms, solve Eq. (16) and substitute into Eq. (11), we find for the scalar conductivity

$$\sigma^+=\tfrac{3}{4}\omega_P^2 iV^*\Bigg[1+4\delta_1/3-\tilde{\omega}\delta_1 V^*-\frac{(1+4\delta_1/3)}{[1+(V^*/\tilde{\omega}W^{*2})(1/(9\delta_2)+4/15-\omega\gamma W^*)]}\Bigg]^{-1}, \tag{20}$$

where $\gamma=(\tilde{\omega}-\omega_c^*)/qV_F^*$, and $\omega_P=(4\pi ne^2/m)^{1/2}$. The functions V^* and W^* are defined by

$$V^*=\int_\pi^0 \frac{\sin^2\theta d(\cos\theta)}{(\tilde{\omega}-\omega_c^*-qV_F^*\cos\theta)}, \tag{21}$$

$$W^*=\int_\pi^0 \frac{\sin^2\theta\cos\theta d(\cos\theta)}{(\tilde{\omega}-\omega_c^*-qV_F^*\cos\theta)}. \tag{22}$$

The conductivity, although it is reasonably complicated, approaches some rather simple and physically meaningful limiting values. In the limit $q\to 0$, σ_+ approaches

$$(\sigma_+)_{q\to 0}=i\omega_P^2/(\tilde{\omega}-\omega_c)\,. \tag{23}$$

There are no effects due to correlations in this limit. This is a consequence of the fact that total momentum is conserved in electron-electron collisions. The induced-current (for a zero wave vector) electric field is not affected by the electron-electron interactions. In the limit $\omega/\omega_c\to 0$, the effect of correlations to lowest order in ν_c/ω_c vanishes. This is true not only for our truncated scattering function, but for a general scattering function as well. If we examine the transport Eq. (3) and the second expression for the current Eq. (12), then neglecting the relaxation term and the time derivative term we see that $(g+E_1)$ satisfies the usual transport equation, without correlations. The energy delta function in Eq. (12), is just sufficient to produce a factor m^* canceling the $1/m^*$ in front of the integral leading to the uncorrelated conductivity.

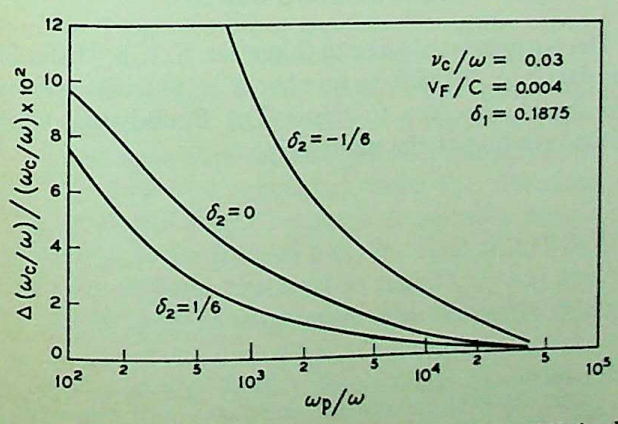


FIG. 3. A plot as a function of ω_p/ω of the relative shift in the position of the Doppler shifted absorption edge due to correlation effects.

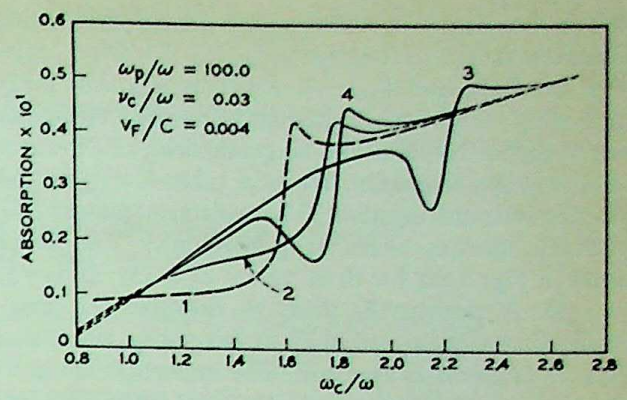


FIG. 4. A plot of the Doppler shifted absorption edge as a function of magnetic field. The plasma frequency, collision frequency and Fermi velocity are fixed. The four curves show the effect of correlation on the edge. The curve labeled (1) has a $\delta_1=0$ and a $\delta_2=0$, i.e., uncorrelated. Curve (2) has $\delta_1=0.1875$, $\delta_2=0$. Curve (3) has $\delta_1=0.1875$, $\delta_2=-0.166$, and curve (4) has $\delta_1=0.1875$, $\delta_2=-0.05$.

III. EVALUATION AND DISCUSSION OF THE SURFACE IMPEDANCE

Typically, the absorption coefficient of a semi-infinite metallic plasma exhibits a rather sharp resonance at the Doppler shifted resonance frequency.[11] If we substitute for q in Eq. (1), the value which is obtained from a simple analysis which leaves out nonlocal effects, i.e.,

$$q^2/q_0^2=-\omega_p^2/[\omega(\omega-\omega_c)]\,. \tag{24}$$

Then we may rewrite the approximate Doppler shifted resonance condition Eq. (1) as

$$\omega_c/\omega=1+[(V_F/c)(\omega_p/\omega)]^{2/3}. \tag{25}$$

At low frequencies, where $(\omega_p/\omega V_F/c)\gg 1$ and $\omega_c/\omega\gg 1$, the effects of correlations are not observable. At higher frequencies, where $\omega_p/\omega V_F/c\sim\omega_c/\omega\sim 1$, correlations are important.

Since the simple resonance condition, Eq. (24), at these "high frequencies" is now a function of the "mass" of the particles, in this case the bare mass, and since correlations do play a role, it seems natural to ask what mass comes into play. There are, of course, two masses in the problem; the so-called bare mass and the clothed mass [see Eq. (7)]. In Fig. 3 we have plotted the

[11] J. Kirsch and P. B. Miller, Phys. Rev. Letters 9, 421 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

shift in frequency of the Doppler edge from its uncorrelated value as a function of (ω_p/ω). For fixed values of $(\nu_c/\omega)=0.03$ and $V_F/c=0.004$ there is a shift away from what might be called a bare-mass resonance toward higher values of the magnetic field, i.e., towards a clothed mass resonance. We have taken $\delta_1=\frac{3}{16}$ so that the effective mass equals 1.25 bare electron masses (the observed effective mass in potassium).[12] The three curves in Fig. 3 are for three values of δ_2 ($\delta_2=0$, $\delta_2=\frac{1}{6}$, $\delta_2=-\frac{1}{6}$). A positive δ_2 tends to decrease the magnitude of the computed shift in frequency whereas a negative δ_2 increases the shift. The important thing to note here is that the resonance occurs at *neither* the frequency determined by the bare mass nor by the clothed mass. *The shift is a dynamical function of the frequency.*

In Fig. 4 we have plotted the actual absorption line for fixed values of the plasma frequency $\omega_p/\omega=100$, Fermi velocity, collision frequency, and δ_1. The effect of a negative δ_2 on the line shape is rather striking. It produces an over-all smearing of the line relative to the uncorrelated line. The magnitude of the steeply rising portion of the absorption curve is severely reduced. In addition, a negative δ_2 causes the absorption edge to exhibit a rather prominent minimum. In some cases (i.e., for the case $\delta_2=-\frac{1}{6}$ or $-\frac{1}{10}$) the minimum or hole in the absorption edge effectively causes a splitting of the edge into two peaks.

Physically, we can understand the origin of this decrease in absorption. Suppose for the purpose of this rather crude argument that the electric field inside the metal could be characterized by a single wave number q, as it can be in the local or classical theory. The edge in the absorption is due to a sharp increase in the conductivity when $\omega_c/\omega\approx(1+qV_F/\omega)$. The electrons at the Fermi velocity traveling in the direction of the wave see a static dc field spiral out around the lines of force and pick up energy from the field. Suppose we now ask ourselves if it is possible to introduce a mechanism which, in this region of increasing conductivity, would tend to decrease the induced current or conductivity, thus producing an associated dip in the absorption edge. Electron interactions can produce just this effect. As the quasiparticles (the current carriers in a Fermi liquid theory) are dragged through the surrounding, now slightly incompressible fluid, they create a backflow by pushing other quasiparticles out of their way. This backflow carries current. At a definite wave number q and frequency ω, it is possible for the backward flowing current to exactly cancel the forward flowing current so that the conductivity goes to zero. For the circularly polarized mode we simply require that

$$\int g(\theta)\sin\theta d(\cos\theta)\sim\Omega_1=0. \tag{26}$$

This implies that

$$1+\frac{V^*}{\tilde{\omega}W^{*2}}\left(\frac{1}{9\delta_2}+\frac{4}{15}-\tilde{\omega}\gamma W^*\right)=0. \tag{27}$$

For a finite δ_2 (second moment) it is possible to find at least one solution to Eq. (27) (to zeroth order in ν_c/ω). The dips in the conductivity, which show up in this simple model are characteristic of the interacting Fermion system. Had we retained more terms in the multipole expansion of the scattering function, it is likely that there would exist multiple solutions of Eq. (27). In the actual boundary value problem, q is not well defined and the zero in the conductivity appears as a minimum in the absorption curve. If the minimum happens to fall in the high field or classical region of the curve, then the magnitude of the effect is reduced since in this limit $q\rightarrow 0$ and there are no effects due to correlations.

The observation of correlation effects in the alkali metals, unfortunately, is extremely difficult since the magnetic fields required in order to keep the Dopper shifted edge in the neighborhood of the cyclotron frequency are enormous. However, experiments could be performed in semimetals, semiconductors, or doped insulators.

One should look for a material, preferably one with a simple band structure, which has an $r_s>1$. In addition, $\omega_p/\omega v_F/c\sim\omega_c/\omega\sim 1$ and $\omega/\nu_c\gg 1$. The frequency range in which one can work will primarily be limited by the magnetic fields which are available.

ACKNOWLEDGMENTS

The authors would like to thank Dr. S. J. Buchsbaum and Dr. P. A. Wolff for a number of helpful discussions. We would also like to thank Dr. Buchsbaum for a critical reading of the manuscript.

[13] C. C. Grimes, Bull. Am. Phys. Soc. (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Low-Field de Haas–van Alphen Effect in Copper*

A. S. Joseph
North American Aviation Science Center, Canoga Park, California
AND
A. C. Thorsen
Atomics International Division of North American Aviation, Inc., Canoga Park, California
(Received 2 January 1964)

The de Haas–van Alphen effect has been observed in a copper single crystal by means of a null-deflection torsion balance technique. Detailed measurements have been carried out which provide an accurate determination of the topology of the Fermi surface necks in copper. Effective mass values are found to be in excellent agreement with cyclotron resonance results. Unusual behavior observed in the angular dependence of the amplitude of the oscillatory torque is attributed to the spin-splitting factor in the theoretical expression for the magnetization.

I. INTRODUCTION

THE first observation of the de Haas–van Alphen effect in copper by Shoenberg was accomplished by means of an impulsive field technique. His detailed results[1] confirmed the model of the Fermi surface proposed by Pippard,[2] and demonstrated conclusively that the Fermi surface deviates from sphericity and is multiply connected along the ⟨111⟩ directions, forming "necks" at the hexagonal faces of the Brillouin zone. The precise size and shape of the necks, however, could not be ascertained with pulsed magnetic fields. In this paper we report on detailed low-field measurements of the de Haas–van Alphen (dHvA) effect in copper resulting in accurate determinations of the neck topography and the associated effective masses. In addition, we report on what we believe to be the first unequivocal observation of the vanishing of the amplitude of the dHvA oscillations due to the zero value of the spin-splitting factor for $m^*=0.5m_0$ (and $g=2$).

II. EXPERIMENTAL

The measurements reported herein were made with a torsion balance apparatus which has been described in detail in a previous publication.[3] Magnetic fields up to 40 kG were supplied by a 22-in. Varian magnet, and were measured to within 0.1% with a rotating coil probe. The magnetic fields were swept automatically at the slow rate of ≈200–400 G/min to reduce the effects of eddy currents in the sample. The copper single crystal, supplied by K. R. Garr of Atomics International, was grown by the Bridgeman technique from a 99.999% pure rod purchased from the American Smelting and Refining Company. A suitable sample was spark cut in the form of a 0.1-in.-diam ×0.15-in.-long cylinder and was mounted so that the [1$\bar{1}$0] axis was along the axis of suspension. The magnetic field could thus be rotated in the (1$\bar{1}$0) plane of the crystal and was oriented in this plane to within 1°.

III. RESULTS AND DISCUSSION

The values of the period of the dHvA oscillations arising from extremal orbits around the necks are shown in Fig. 1. The period has a maximum value of 4.64×10^{-8} G^{-1} along the [111] axis (≈12% greater than that reported in Ref. 1) corresponding to a minimum cross-sectional area of 2.06×10^{15} cm^{-2}. As the field is rotated away from the [111] direction the period decreases and can be determined over an angular range of approximately 48°. To within the experimental accuracy, the data can be represented by the expression

$$\frac{P(\theta)}{P(0)}=\frac{m^*(0)}{m^*(\theta)}=\cos\theta\left(1-\frac{m_t}{m_l}\tan^2\theta\right)^{1/2}, \qquad (1)$$

with $m_t/m_l=2.55\pm0.10$, indicating that the shape of

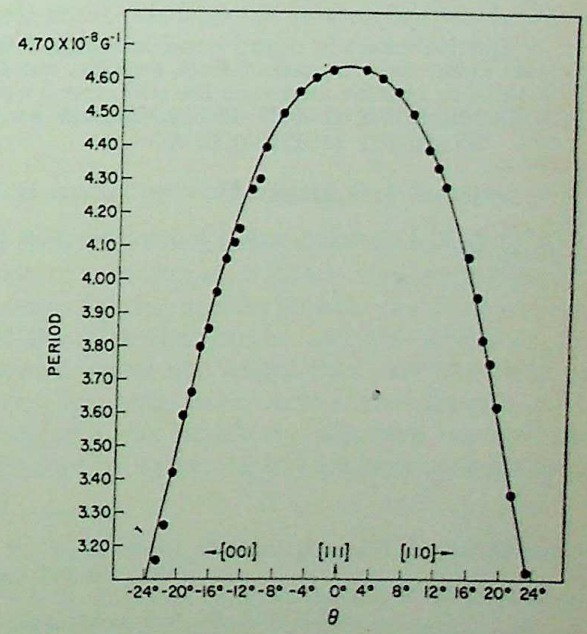


FIG. 1. de Haas-van Alphen period in copper versus orientation in the (1$\bar{1}$0) plane. The solid line is a plot of Eq. (1) with the parameter $m_t/m_l=2.55$.

* Supported in part by the U. S. Atomic Energy Commission.
[1] D. Shoenberg, Phil. Trans. Roy. Soc. London A255, 85 (1962).
[2] A. B. Pippard, Phil. Trans. Roy. Soc. London A250, 325 (1957).
[3] A. S. Joseph and W. L. Gordon, Phys. Rev. 126, 489 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the neck closely approximates the form of a hyperboloid of one sheet, and is hence similar to that of nickel.[4] In the above equation, θ is the angle measured from the [111] axis in the $(1\bar{1}0)$ plane, and m_t and m_l are the transverse and longitudinal effective masses.

At several angles the effective mass was determined from the temperature dependence of the amplitude of the oscillations. These results are shown in Fig. 2 and are compared with Eq. (1), indicated by the solid line. The theoretical curve was normalized to the value $m^*(13°)=0.5m_0$ at $\theta=\pm 13°$ for reasons to be discussed below. The data indicate a value of effective mass along the [111] axis of $0.45m_0\pm 3\%$. These results are in excellent agreement with the masses determined by Koch, Stradling, and Kip[5] by means of cyclotron resonance.

In Fig. 3 the amplitude of the oscillatory torque is plotted as a function of θ for a constant magnetic field. At $\theta=0$, this amplitude vanishes as is always the case for an extremum in the magnitude of the period. The torque amplitude increases initially with θ, passes through a maximum and goes to zero at $\theta=\pm 13°\pm 1°$. For larger angles the amplitude again increases and then decreases, finally vanishing at $\theta\approx\pm 24°$. We attribute the vanishing of the amplitude at $\theta\approx\pm 13°$ to the effect of spin-splitting of the Landau levels. This effect is manifested in a cosine multiplicative factor in the theoretical expression for the magnetization. The factor can be expressed as $\cos[\pi g m^*/2m_0]$, where g is the spectroscopic splitting factor, and m^* is the effective mass. It was introduced by Sondheimer and Wilson[6] and Dingle[7] for the case when m^* differs from m_0, and is given above in the generalized form derived by Cohen and Blount.[8] If the reasonable assumption that $g=2$ is made, then the amplitude should vanish when $m^*=0.5\ m_0$. Our experimental mass values do indeed approach this value for $\theta\approx\pm 13°$, evidencing remarkable agreement with the theory.

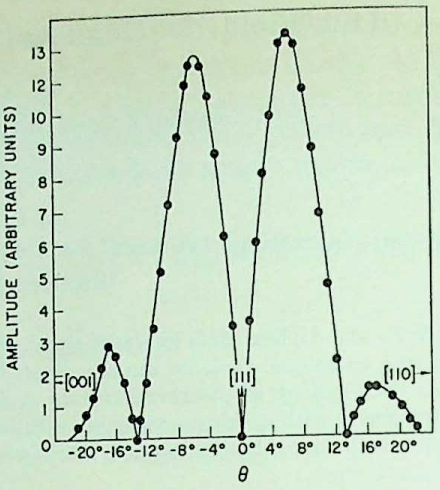


FIG. 3. Amplitude of the oscillatory torque at constant magnetic field (in arbitrary units) versus orientation in the $(1\bar{1}0)$ plane.

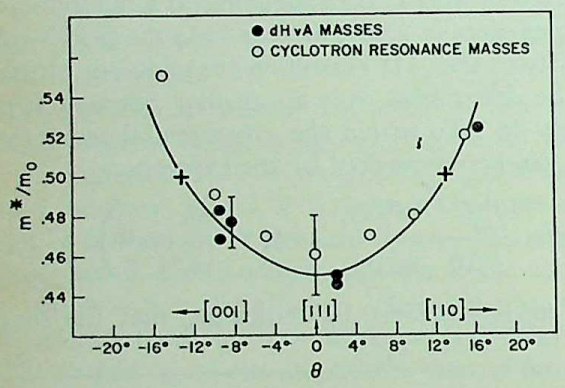


FIG. 2. Effective masses in copper versus orientation in the $(1\bar{1}0)$ plane. Open circles are data of Koch, Stradling, and Kip (Ref. 5), found by cyclotron resonance. The solid curve is a plot of Eq. (1) normalized to a value $m^*=0.5\ m_0$ at the points denoted by crosses.

ACKNOWLEDGMENTS

The authors wish to thank R. E. Gile for his assistance in taking the data and D. G. Swarthout for orienting the crystal. We are indebted to K. R. Garr for supplying the high-purity single crystal of copper used in this study and to J. K. Koch, R. A. Stradling, and A. F. Kip for communicating their cyclotron resonance data prior to publication.

[4] A. S. Joseph and A. C. Thorsen, Phys. Rev. Letters **11**, 554 (1963).
[5] J. F. Koch, R. A. Stradling, and A. F. Kip, Phys. Rev. **133**, A240 (1964).
[6] E. H. Sondheimer and A. H. Wilson, Proc. Roy. Soc. (London) **A210**, 173 (1951).
[7] R. B. Dingle, Proc. Roy. Soc. (London) **A211**, 500 (1952).
[8] M. Cohen and E. I. Blount, Phil. Mag. **5**, 115 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Lattice Dynamics and Phase Transitions of Strontium Titanate

R. A. COWLEY*
Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England
(Received 16 December 1963)

The crystal dynamics of strontium titanate has been studied both experimentally and theoretically. The frequency versus wave-vector dispersion curves for some of the normal modes propagating along the [0,0,1] direction have been measured by neutron spectrometry at 90 and 296°K. The experiments were performed at the Chalk River Laboratories of Atomic Energy of Canada Ltd., using a triple-axis crystal spectrometer. The temperature dependence of the transverse optic mode of the lowest frequency has been found to be in agreement with the temperature dependence of the dielectric constant, as predicted by Cochran. The experimental results have been used to obtain the parameters of several models, more than one of which gives reasonable agreement with the experimental results. It is suggested that the anomalous behavior of the elastic properties and the phase transition at 110°K are associated with an accidental degeneracy of two branches of the dispersion curves; the longitudinal acoustic branch and the transverse optic branch of lowest frequency. The origin of the temperature dependence of this transverse optic mode and the relevance of lattice dynamics to the phase transitions in other perovskites are discussed.

I. INTRODUCTION

THE lattice dynamics of strontium titanate has been studied both theoretically and experimentally, in an attempt to understand the lattice dynamics of a ferroelectric material. Anderson[1] and Cochran[2] have independently suggested that the anomalous temperature dependence of the static dielectric constant of ferroelectrics is associated with the temperature dependence of a transverse optic mode of vibration. Recently, this low-frequency mode of vibration has been observed experimentally in barium and strontium titanates, by using infrared spectrometry[3,4] and in strontium titanate by using neutron spectrometry.[5]

In this paper experimental measurements are presented of some of the frequency versus wave-vector dispersion curves for normal modes propagating in the [0,0,1] direction of strontium titanate. These dispersion curves were determined by neutron spectrometry using the triple-axis crystal spectrometer[6] at the N.R.U. reactor at the Chalk River Laboratories of Atomic Energy of Canada Limited.

The simplest model which is used to describe the lattice dynamics of ionic crystals is the rigid-ion model.[7,8] In this model the ions interact with one another through both long-range electrostatic forces and short-range repulsive forces between neighboring ions. This model has been shown to be inadequate to describe the measured dispersion curves of two alkali halides, sodium iodide[9,10] and potassium bromide.[10] These experimental results showed that a more satisfactory model was the shell model,[9] which takes account of the polarizability of the ions in both electrostatic and short-range forces.

The experimental measurements on strontium titanate have been used to obtain the parameters of both rigid-ion models and shell models. These parameters show that it is possible to obtain the large changes in the frequency of the lowest transverse optic mode by changing the values of the parameters only slightly.

Recently, it has been found that the ultrasonic attenuation in strontium titanate increases dramatically below 110°K, while the elastic constants show an anomalous temperature dependence.[11,12] A possible explanation for these observations is put forward in terms of an instability of the crystal arising from the accidental degeneracy of two branches of the dispersion curves. Finally the results are discussed with particular reference to the origin of ferroelectricity in the perovskite structure.

II. THEORY OF THE LATTICE DYNAMICS OF STRONTIUM TITANATE

1. Symmetry

Some of the properties of the normal modes of vibration of a crystal are a direct consequence of its symmetry. The degeneracies between the different branches of the dispersion curves, and the separation of the normal modes into longitudinal and transverse vibrations for some wave vectors are examples of these properties. For simple crystals these relations can be obtained by inspection, but for more complicated crys-

* Visitor at Atomic Energy of Canada Limited, 1961–62, where some of the work described below was done.

[1] P. W. Anderson, Izd.AN. SSR., Moscow (1960).
[2] W. Cochran, Advan. Phys. 9, 387 (1960).
[3] A. S. Barker and M. Tinkham, Phys. Rev. 125, 1527 (1962).
[4] W. G. Spitzer, R. C. Miller, D. A. Kleinman, and L. E. Howarth, Phys. Rev. 126, 1710 (1960).
[5] R. A. Cowley, Phys. Rev. Letters 9, 159 (1962).
[6] B. N. Brockhouse, *Inelastic Scattering of Neutrons in Solids and Liquids* (International Atomic Energy Agency, Vienna, 1961), p. 113.
[7] E. W. Kellermann, Phil. Trans. Roy. Soc. A238, 513 (1940).
[8] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, London, 1954).
[9] A. D. B. Woods, W. Cochran, and B. N. Brockhouse, Phys. Rev. 119, 980 (1960).
[10] A. D. B. Woods, B. N. Brockhouse, R. A. Cowley, and W. Cochran, Phys. Rev. 131, 1025 (1963).
[11] R. O. Bell and G. Rupprecht, Phys. Rev. 129, 90 (1963).
[12] K. S. Krogstad and R. W. Moss, Bull. Am. Phys. Soc. 7, 192 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

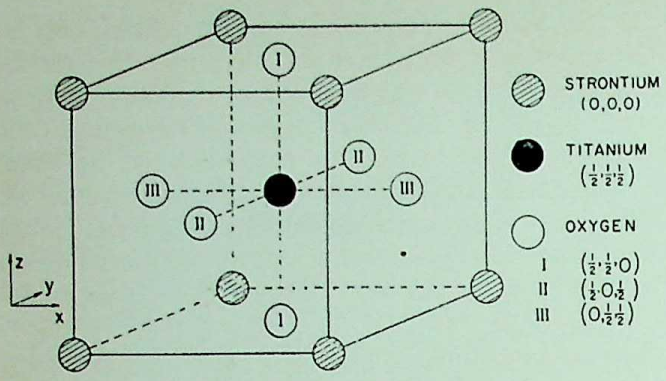


FIG. 1. The crystal structure of strontium titanate.

tals a more systematic procedure is given by using group theory.[13]

The normal modes of vibration are classified by assigning them to an irreducible representation of the space group of the crystal. A representation for the normal modes is obtained by choosing as a set of basis vectors the displacement of each ion in turn along a coordinate axis. If the character of a symmetry operation S is $\chi(S)$ in this representation, and its character in the λ irreducible representation is $\chi^\lambda(S)$, then

$$\chi(S)=\sum_\lambda C_\lambda \chi^\lambda(S), \qquad (1)$$

where C_λ is the number of times the λ irreducible representation occurs in the original representation.

When the normal modes of the crystal are described by Bloch waves, the irreducible representations of the symmorphic space groups can be obtained from the irreducible representations of the little group, which is that point group which leaves the wave vector unchanged. The irreducible representations are then obtained from Eq. (1) and from the characters of the irreducible representations of the point groups.[13,14]

The structure of strontium titanate is cubic perovskite, and the space group is $Pm3m$. The ions are situated on five interpenetrating simple cubic lattices as shown and labeled in Fig. 1. The characters of the different symmetry elements in the original representation are listed in Appendix I.

(a) $\zeta=(0,0,0)$

The little group is $m3m$, and the irreducible representations are

$$4\Gamma_{15}+\Gamma_{25}.$$

This result is, however, incorrect because the effect of the macroscopic electric field has been neglected. This electric field splits the degeneracy of some of the normal modes because the boundary conditions for the longitudinal and transverse modes differ.[8] The degeneracies will be lifted for those normal modes which transform like an ordinary vector. Since an ordinary vector transforms like Γ_{15}, the degeneracies associated with three of the four irreducible representations Γ_{15} are split. The fourth irreducible representation corresponds to a uniform translation of the whole crystal. There are, therefore, the three acoustic modes, $\omega=0$, a triply degenerate normal mode Γ_{25}, and three longitudinal and doubly degenerate transverse optical modes.

(b) $\zeta=(0,0,\zeta)$

The little group is $4mm$, and the irreducible representations are

$$4\Delta_1+\Delta_2+5\Delta_5.$$

The normal modes Δ_1 and Δ_2 are longitudinally polarized modes, while those represented by Δ_5 are doubly degenerate transverse modes.

At the zone boundary the little group is $4/mmm$, and the irreducible representations are

$$2M_1+2M_2'+M_3+3M_5+2M_5'.$$

M_5 and M_5' correspond to doubly degenerate transverse modes, while the other modes are longitudinally polarized.

(c) $\zeta=(\zeta,\zeta,0)$

The little group is mm except at the zone boundary, where it is $4/mmm$. The irreducible representations are

$$5\Sigma_1+\Sigma_2+5\Sigma_3+4\Sigma_4,$$

and

$$M_1+M_2+M_2'+M_3+2M_3'+M_4+M_5+3M_5'.$$

There are no degeneracies at general points, but at the zone boundary there are four doubly degenerate normal modes.

(d) $\zeta=(\zeta,\zeta,\zeta)$

The little group at a general point is $3m$ and its irreducible representations are

$$4\Lambda_1+\Lambda_2+5\Lambda_3.$$

At the zone boundary the little group symmetry is increased to $m3m$ and the irreducible representations are

$$\Gamma_2'+\Gamma_{12}'+\Gamma_{25}+\Gamma_{25}'+2\Gamma_{15}.$$

The irreducible representation Λ_3 corresponds to doubly degenerate normal modes, while at the zone boundary there are four triply degenerate normal modes, one doubly degenerate pair and a single normal mode.

2. Models

In the last section we found the properties of the normal modes which can be deduced from the symmetry of the crystal. In order to evaluate the frequencies of any of the normal modes, a particular model must be

[13] V. Heine, *Group Theory in Quantum Mechanics* (Pergamon Press, Inc., New York, 1960).
[14] G. F. Koster, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. V.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

employed for the interactions between the ions. The simplest model which has been used for ionic crystals is the rigid-ion model.[7,8] The equation of motion for this model is then written in the harmonic approximation as

$$M_K\omega^2\mathbf{U}(K)=\sum_{K'}(\mathbf{R}_{KK'}+Z_K\mathbf{C}_{KK'}Z_{K'})\mathbf{U}(K'), \quad (2)$$

where the notation is the same as that of Woods *et al.*[9]

This model does not satisfactorily account for the dispersion curves of the alkali halides,[9,10] because it neglects the polarizabilities of the ions. These polarizabilities can be introduced by using the shell model.[2,9] The equations of motion of the shell model in the harmonic approximation[9] are

$$\begin{aligned} M_K\omega^2\mathbf{U}(K)&=\sum_{K'}(\mathbf{R}_{KK'}+Z_K\mathbf{C}_{KK'}Z_{K'})\mathbf{U}(K') \\ &\quad+\sum_{K'}(\mathbf{T}_{KK'}+Z_K\mathbf{C}_{KK'}Y_{K'})\mathbf{W}(K'), \\ 0&=\sum_{K'}(\mathbf{T}_{K'K}+Y_K\mathbf{C}_{KK'}Z_{K'})\mathbf{U}(K') \\ &\quad+\sum_{K'}(\mathbf{S}_{KK'}+Y_K\mathbf{C}_{KK'}Y_{K'})\mathbf{W}(K'). \end{aligned} \quad (3)$$

The matrices $\mathbf{T}_{KK'}$ and $\mathbf{S}_{KK'}$ give the short-range interactions between the displacements and the electronic dipole moments, and between the electronic dipole moments, respectively. In general, the elements $\mathbf{R}_{KK'}$, $\mathbf{T}_{KK'}$ and $\mathbf{S}_{KK'}$ are all different and must be specified by different short-range force constants. However the shell charges Y_K can be chosen so that $\mathbf{R}_{KK'}=\mathbf{T}_{KK'}$ when the wave vector is zero,[15] and it was further assumed that they are then equal for all wave vectors. The validity of this approximation and the further approximation that $\mathbf{S}_{KK'}=\mathbf{R}_{KK'}+\mathbf{k}_K\delta_{KK'}$ are discussed by Cowley, Cochran, Woods, and Brockhouse.[15] The model then corresponds to a shell model in which all the short-range forces act through the shells.

In actual calculations the short-range forces were assumed to be axially symmetric. If the equilibrium condition is imposed then these forces become central forces and the elastic constants obey Cauchy's relation, $C_{12}=C_{44}$. In practice, this relation is nearly satisfied, but the equilibrium condition was not imposed on the models described below.

For each type of short-range interaction two parameters are needed to specify axially symmetric forces. These parameters were chosen to be the derivatives of a potential function parallel and perpendicular to the line joining the interacting ions. These were defined in a similar way to those used for the alkali halides.[7,9,15]

The forces were specified by

$$\left(\frac{\partial^2V}{\partial r^2}\right)_{\parallel}=\frac{e^2A_i}{2v}, \quad \left(\frac{\partial^2V}{\partial r^2}\right)_{\perp}=\frac{e^2B_i}{2v},$$

where the suffix i is one for strontium-oxygen, two for titanium-oxygen, and three for oxygen-oxygen forces. Each type of force was assumed to act only between nearest neighbors. The short-range contributions to the dynamical matrix were then evaluated by the usual procedure,[8] and the elements are listed in Appendix II.

The ionic charges were specified by two charges; that on the strontium ion was Z_1e and on the titanium ion Z_2e. The charge on the oxygen ion is then $Z_3e=-\frac{1}{3}(Z_1+Z_2)e$.

The polarizabilities of the ions were specified in a similar way to that used for the alkali halides.[9,15] These polarizability parameters were chosen so that they were independent of the choice of the shell charges. The electrical polarizability was defined as

$$\alpha_i=Y_i^2/[k_i+(T_{ii})_0],$$

and the short-range polarizability as

$$d_i=-Y_i(T_{ii})_0/[k_i+(T_{ii})_0],$$

where the suffix i has values 1 for the strontium ion, 2 for the titanium and 3 for the oxygen, and $(T_{11})_0=2A_1+4B_1$, $(T_{22})_0=A_2+2B_2$. The oxygen-ion polarizabilities are not isotropic, so that the polarizability was defined in terms of a mean $(T_{33})_0$ given by,

$$(T_{33})_0=\tfrac{1}{3}(2A_1+4B_1+A_2+2B_2+4A_3+8B_3).$$

The most complicated rigid-ion model of the crystal, then, has eight adjustable parameters, six for the short-range forces and two for the ionic charges, while for a shell model there are fourteen parameters, the same eight and six polarizability parameters.

3. Calculation of Dielectric and Elastic Constants

The frequencies of the normal modes of vibration can be obtained by solving Eq. (2) for a rigid-ion model, or Eqs. (3) for a shell model. Once a particular set of parameters has been chosen, their solution is quite straightforward when the electrostatic coupling coefficients are known. These coefficients were calculated using the EDSAC II computer at the Cambridge University Mathematical Laboratory, and their values for the (0,0,1) direction have already been reported.[16] These and the coefficients for the (1,1,0) and (1,1,1) directions are listed in Appendix III.

The dielectric constants can be evaluated in terms of the parameters introduced above.[2,17] The high-frequency dielectric constant is given by Eq. (2.44) of Ref. 17 and the static dielectric constant by Eq. (2.34). In practice the high-frequency dielectric constant was evaluated explicitly and the static dielectric constant obtained from the Lyddane-Sachs-Teller relation as extended by Cochran.[2]

[15] R. A. Cowley, W. Cochran, A. D. B. Woods, and B. N. Brockhouse, Phys. Rev. 131, 1030 (1963).

[16] R. A. Cowley, Acta Cryst. 15, 687 (1962).

[17] R. A. Cowley, Proc. Roy. Soc. (London) A268, 121 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The elastic constants were evaluated explicitly using the method of long waves as developed by Born and Huang[8] and Cowley.[17] Since every ion is situated at a center of symmetry, the polarizabilities do not affect the elastic constants. The electrostatic contributions were evaluated numerically.[16] The elastic constants are given by:

$$C_{11}=(2e^2/vr)[\tfrac{1}{8}(2A_1+2B_1+A_2+2A_3+2B_3)+0.16485(Z_1^2+Z_2^2)-0.78348Z_3^2-1.0721Z_1Z_2+2.67981Z_2Z_3-1.27803Z_1Z_3],$$

$$C_{12}=(2e^2/vr)[\tfrac{1}{8}(A_1-5B_1-B_2+A_3-5B_3)-0.55532(Z_1^2+Z_2^2)-1.60946Z_3^2+0.26873Z_1Z_2-1.43585Z_2Z_3+0.05648Z_1Z_3],$$

$$C_{44}=(2e^2/vr)[\tfrac{1}{8}(A_1+3B_1+B_2+A_3+3B_3)-0.08242(Z_1^2+Z_2^2)+0.39174Z_3^2+0.53605Z_1Z_2-1.33991Z_2Z_3+0.63902Z_1Z_3].$$

The electrostatic cohesive energy is $(e^2/r)\alpha_M$, where α_M, the Madelung constant, is given by

$$1.25950(Z_1^2+Z_2^2)+4.57109Z_3^2+0.79260Z_1Z_3-0.667170Z_2Z_3+0.48362Z_1Z_2.$$

The equilibrium condition is then

$$4\alpha_M+12(B_1+B_3)+3B_2=0.$$

This last equation when substituted into the equation for C_{12} gives $C_{12}=C_{44}$. r is half the lattice parameter and v the volume of the unit cell, $v=8r^3$.

III. THE EXPERIMENTS

1. Experimental Methods

The excellent single crystal of strontium titanate was a 150-carat boule provided by the Titanium Division of the National Lead Company. The mosaic spread was measured against that of a silicon crystal and was 0.25 deg. The crystal, which had [1,0,0] aligned along the length of the boule, was mounted in a metal cryostat, and could be either cooled or heated from above. It was aligned *in situ* by using neutrons.

The experiments were carried out by using a triple-axis crystal spectrometer.[6]

The integrated intensity of a neutron group, obtained using the "constant Q" technique,[6] is given in the harmonic approximation by[9]

$$\sigma=\frac{\hbar}{8\pi^2}\frac{|\mathbf{k}'|}{|\mathbf{k}_0|}\frac{H(\mathbf{q}j)^2}{\omega(\mathbf{q}j)}\left\{\begin{matrix}N(\mathbf{q}j)\\N(\mathbf{q}j)+1\end{matrix}\right\}.$$

$H(\mathbf{q}j)$ is the structure factor for the mode and is given by

$$H(\mathbf{q}j)=|\sum_K \exp(2\pi i\boldsymbol{\tau}\cdot\mathbf{R}(0K))\mathbf{Q}\cdot\mathbf{e}(K,\mathbf{q}j)b_K M_K^{-1/2}\times\exp(-W(K))|.$$

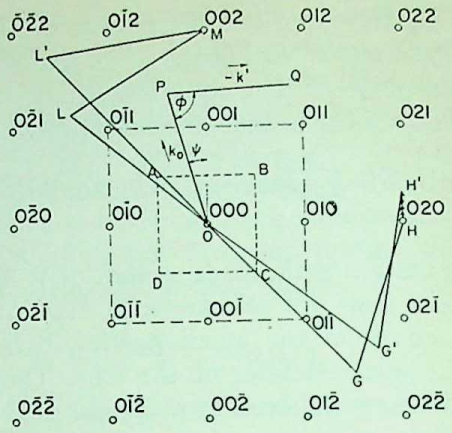


FIG. 2. The (1,0,0) plane of the reciprocal lattice of strontium titanate. The Brillouin zone is labeled ABCD and the area in which the reduced structure factor repeats is enclosed by (0,1,1), (0,1,$\bar{1}$), (0,$\bar{1}$,$\bar{1}$) and (0,$\bar{1}$,1). OLM and OL'M demonstrate the "constant-Q" technique in which the spectrometer is controlled so as to keep the momentum transfer constant while varying the energy transfer, while OGH and OG'H' shows the constant-energy technique in which the momentum is kept along the desired direction.

If a normal mode is propagating along the [0,0,1] direction then the motion of all the ions is either transverse or longitudinal. The dependence of the structure factor on the wave-vector transfer can then be separated out if we assume the Debye-Waller factors for the different atoms to be equal. We can then introduce a reduced structure factor for the normal mode

$$h(\mathbf{q}j)=|\sum_K \exp(2\pi i\boldsymbol{\tau}\cdot\mathbf{R}(0K))e(K,\mathbf{q}j)M_K^{-1/2}b_K|.$$

This reduced structure factor is periodic in reciprocal space. Figure 2 shows the (1,0,0) plane of the reciprocal lattice of strontium titanate. The Brillouin zone is labeled as ABCD, while the reduced-structure factor repeats over a larger unit in reciprocal space, enclosed on Fig. 2 by the reciprocal lattice points (0,$\bar{1}$,1), (0,1,1), (0,1,$\bar{1}$), and (0,$\bar{1}$,$\bar{1}$). Because the structure factor repeats over a larger volume in reciprocal space than the frequencies, there are several different points in reciprocal space at which phonons of the same wave vector have different reduced-structure factors. In the (1,0,0) plane there are four corresponding to reciprocal-lattice vectors, $\boldsymbol{\tau}$: (0,0,0), (0,1,0), (0,0,1) and (0,1,1).

The reduced-structure factor squared has been calculated for each of these different positions for normal modes propagating in the (0,0,1) direction, by using model IV of Sec. IV. The results are shown in Fig. 3.

2. Measurements of the Dispersion Curves

Measurements have been made at 90°K of four of the transverse branches and of most of three of the longitudinal branches. The measured branches are those with the lower frequencies. The results are illustrated in Figs. 7 and 9, where they are compared with several of the models described in the next section. The highest frequency transverse optic branch and the two highest

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

frequency longitudinal optic branches were not measured because of the extra difficulty in measuring these high frequencies. Unfortunately it was not possible to complete the measurements on the longitudinal acoustic branch with the crystal in the (1,0,0) plane.

Less complete measurements were made of three of the transverse branches and the longitudinal acoustic branch at 296°K. The results are shown in Figs. 7 and 9, where they are compared with calculations using several of the models described in the next section.

The frequencies of the normal modes with long wavelength, $\mathbf{q}=0$, can also be obtained by using infrared spectroscopy. Strontium titanate has been studied by Last,[18] Barker and Tinkham[3] and by Spitzer *et al.*[4] The frequencies of the transverse modes can be obtained directly from these measurements, and those of the longitudinal modes by using a result due to Kurosawa,[19] as explained by Cochran and Cowley.[20] In Table I the results of the infrared and neutron techniques are listed, for the $\mathbf{q}=0$ modes.

Since in strontium titanate each ion is situated at a center of symmetry, its Raman spectrum is a second-order spectrum and depends upon the joint density of states for the phonons. Narayanan and Vedam[21] have erroneously interpreted their results as a first-order spectrum.

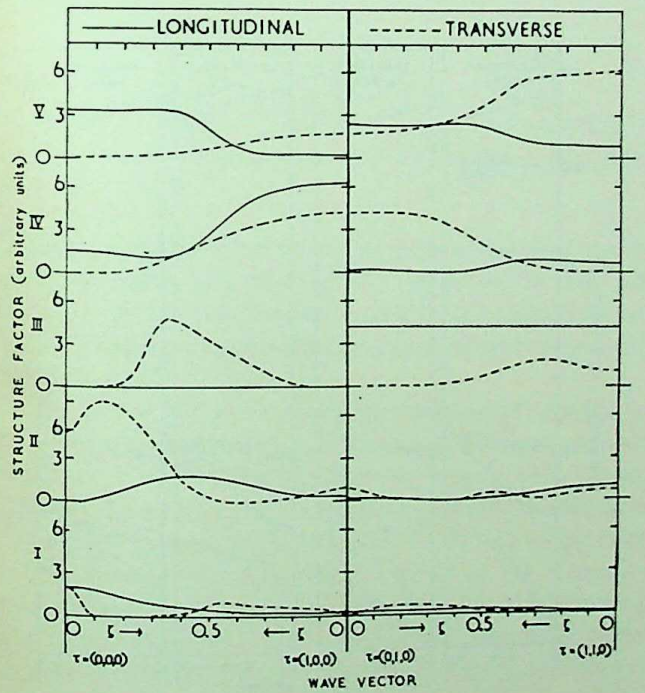


FIG. 3. The reduced structure factor for normal modes propagating along the $(\zeta,0,0)$ direction calculated using model IV of Sec. IV. The solid lines show longitudinal modes and the transverse modes are polarized along the y direction and shown by dotted lines. The modes are labeled in order of increasing frequency.

[18] J. T. Last, Phys. Rev. **105**, 1740 (1957).
[19] T. Kurosawa, J. Phys. Soc. Japan **16**, 1298 (1961).
[20] W. Cochran and R. A. Cowley, *Inelastic Scattering of Slow Neutrons* (International Atomic Energy Agency, Vienna, 1963), Vol. I.
[21] P. S. Narayanan and K. Vedam, Z. Physik **163**, 158 (1961).

TABLE I. The frequencies of the $\mathbf{q}=0$ optical modes in strontium titanate (units; 10^{12} c/sec) as determined by neutron and infrared spectrometry. The measurements with an asterisk are at 90°K and the others at 296°K.

Normal mode	Neutron spectroscopy measurements	Infrared measurements (3)	(4)
Transverse			
ν_1	1.27*	1.20*	...
	2.73	3.00	2.63
ν_2	5.10*	...	5.34
ν_3	7.95*	...	...
ν_4	...	16.5	16.3
Longitudinal			
ν_1	5.10*	...	5.25
ν_2	7.95*	...	...
ν_3	...	14.5	13.8
ν_4	...	24.5	24.9

3. The Temperature Dependence of the Lowest Frequency Transverse Optic Mode

Cochran[2] has shown that, if the static dielectric constant of a crystal follows a Curie law temperature dependence,

$$\epsilon^0=C/(T-T_c), \tag{4}$$

the temperature dependence of the $\mathbf{q}=0$ transverse optic mode with lowest frequency is expected to be

$$\omega_T^2=K(T-T_c). \tag{5}$$

Mitsui and Westphal[22] and Weaver[23] have shown that the static dielectric constant above 70°K follows a Curie Law with a Curie temperature of about 35°K.

The $\mathbf{q}=0$ mode of the transverse optic branch with lowest frequency has been studied at five different temperatures, and the neutron groups obtained at three of these temperatures are shown in Fig. 4. The increase in intensity on the low-frequency side of the 90°K group is due to the elastic and acoustic mode scattering associated with the (0,2,0) reciprocal lattice point. The asymmetry of some of the groups and the change in the width of the neutron groups with temperature are probably largely due to effects associated with the finite resolution of the spectrometer.

Figure 5 shows a plot of the frequencies squared against temperature for this normal mode. The excellent straight line gives ample evidence for Eq. (5), and the validity of discussing ferroelectrics as a problem in lattice dynamics, as suggested by Cochran.[2] The Curie temperature is given by the intercept of the straight line with the temperature axis, which yields a Curie temperature of 32 ± 5°K, which is in excellent agreement with the dielectric constant measurements.[22,23]

[22] T. Mitsui and W. B. Westphal, Phys. Rev. **124**, 354 (1961).
[23] H. E. Weaver, Phys. Chem. Solids **11**, 274 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

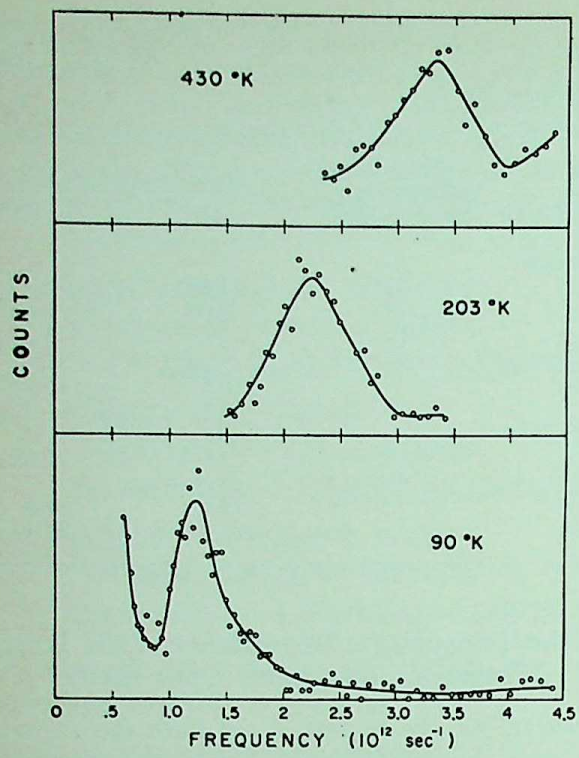


FIG. 4. Neutron groups for the q=0 lowest frequency transverse optic mode at three different temperatures.

4. The Temperature Dependence of the Transverse Acoustic Branch

Recent measurements of the elastic constants of strontium titanate[11,12] have shown that on cooling from room temperature, C_{44} at first increases, but near 110°K it decreases quite dramatically.[11]

Neutron measurements have been made of the $\zeta=(0,0,0.2)$ transverse acoustic mode to see if a similar temperature dependence is found at these higher frequencies. The results are shown in Fig. 6 and there is clearly no evidence for any discontinuity at 110°K, within the limits of experimental error.

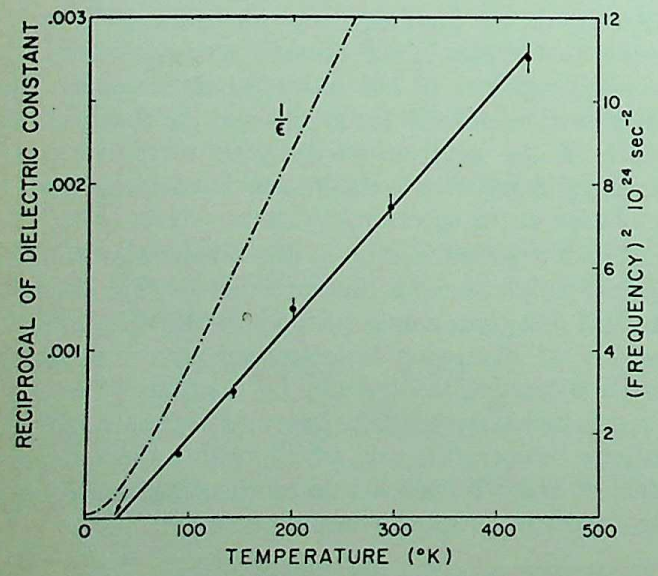


FIG. 5. A plot of the frequency squared against temperature for the q=0 lowest frequency transverse optic mode. The dotted line shows the reciprocal of the static dielectric constant (Refs. 22 and 23).

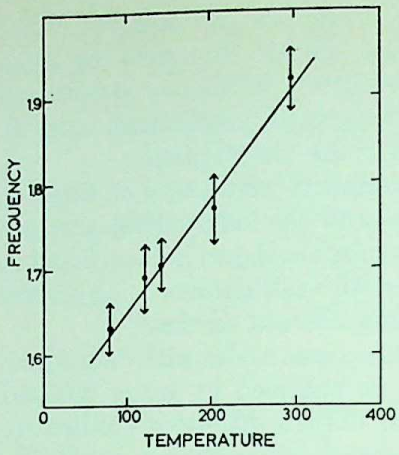


FIG. 6. A plot of the frequency (10^{12} c/sec) against temperature (°K) for the $\zeta=(0.2,0,0)$ transverse acoustic mode.

The ultrasonic measurements also showed anomalous attenuation below 110°K, but no anomalous broadening of the neutron groups at 90°K was detected for either the longitudinal or transverse acoustic branches.

IV. MODELS

The theory of the lattice dynamics of strontium titanate was described in Sec. II. In this section we discuss the results obtained when the parameters of these models are fitted to the experimentally measured elastic[24] and dielectric constants,[22,23,25] and to the fre-

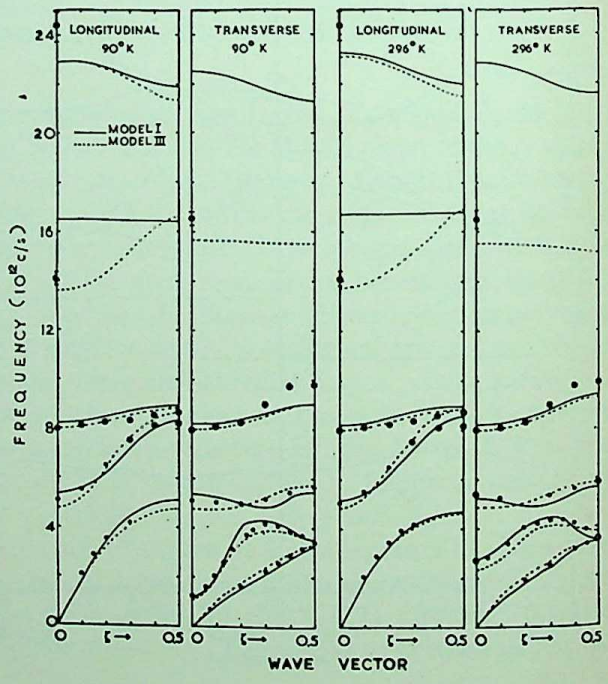


FIG. 7. The dispersion curves for model I (solid line) and model III (dotted line) for the normal modes propagating along the $(\zeta,0,0)$ direction. The experimental measurements are taken from both neutron and infrared results. (See Refs. 3 and 4.) The irreducible representations at the zone boundary are for the longitudinal modes M_2', M_1, M_3, M_1, M_2' in order of increasing frequency, while for the transverse modes they are M_5, M_5', M_5, M_5', M_5 (model I) and M_5', M_5, M_5, M_5', M_5 (model III).

[24] E. Poindexter and A. A. Giardini, Phys. Rev. **110**, 1069 (1958).

[25] S. B. Levin, N. J. Field, F. M. Plock, and L. Merker, J. Opt. Soc. Am. **45**, 737 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. The parameters of the models described in Sec. IV. Those parameters which were varied in each model are given with the errors of the parameters, while for those which were kept fixed there are not any errors.

Model	Short-range forces $e^2/2v$ Strontium-Oxygen A_1	B_1	Titanium-Oxygen A_2	B_2	Oxygen-Oxygen A_3	B_3	Ionic charge (e) Strontium Z_1	Titanium Z_2
90°K								
I	9.0±4.7	2.5±2.1	110.4±18.7	−36.6± 5.7	24.8±6.9	1.0±1.5	1.40±0.21	2.20±0.36
II	26.5±2.1	−2.0±1.3	288.5±10.3	−42.7± 2.4	−3.7±4.3	−0.1±0.9	2	4
III	26.9±2.5	−3.8±1.3	292.2±10.7	−41.4± 5.4	−6.8±5.0	0.8±1.2	2	4
IV	12.9±7.4	0.3±1.5	313.7±19.7	−64.0±18.0	5.3±8.8	0.4±3.9	1.26±0.47	4.64±0.16
V	6.9±1.5	2.6±0.9	321.4± 9.8	−70.1± 4.8	8.8±2.9	−2.1±0.7	0.83±0.07	4.94±0.08
VI	6.5±0.7	2.8±0.4	317.7± 2.4	−68.8± 0.7	9.8±0.6	−2.5±0.3	0.82625	4.9123
296°K								
I	2.9±2.3	6.7±4.6	112.1±20.9	−36.3± 6.5	28.3±7.7	0±1.8	1.20±0.16	2.32±0.35
II	27.6±2.8	−2.8±1.7	289.6±12.9	−39.2± 4.2	−2.9±5.4	−1.1±1.4	2	4
III	28.7±3.2	−4.2±2.0	293.3±12.0	−40.8± 9.2	−5.5±6.2	0.1±1.4	2	4
IV	16.5±7.0	−1.4±2.3	300.5±36.2	−46.1±14.8	5.5±8.6	−1.6±2.8	1.34±0.50	4.39±0.50
V	8.6±2.1	2.1±1.3	315.0±12.5	−70.2± 5.3	9.1±3.9	−2.3±0.9	0.83±0.07	4.91±0.09

Model	Polarizability (10^{-24}) cm³ or (e) Strontium ion α_1	d_1	Titanium ion α_2	d_2	Oxygen ion α_3	d_3	Expected error of an observable/ exptl. error
90°K							
I	0	0	0	0	0	0	4.1
II	0.025	0	0.003	0	0.028±0.002	0.69±0.11	2.6
III	0.025	−0.35±0.29	0.003	−0.08±0.15	0.030±0.002	0.68±0.13	2.5
IV	0.025	−0.29±0.59	0.003	0.01±0.06	0.027±0.004	0.82±0.11	2.3
V	0.025	0	0.068±0.011	−2.13±0.46	0.018±0.003	0.51±0.14	1.7
VI	0.025	0	0.79441	−2.596	0.01535	0.42162	1.6
296°K							
I	0	0	0	0	0	0	4.4
II	0.025	0	0.003	0	0.028±0.004	0.68±0.15	3.3
III	0.025	−0.48±0.23	0.003	−0.05±0.16	0.029±0.003	0.69±0.15	2.8
IV	0.025	−0.18±0.53	0.003	0.03±0.09	0.027±0.003	0.78±0.16	2.7
V	0.025	0	0.079±0.011	−2.60±0.41	0.015±0.005	0.42±0.17	1.9

quencies of the normal modes as measured by infrared[3,4] and neutron spectrometry, as described in the last section. The fitting was performed by a nonlinear least-squares analysis and the numerical work was done using the EDSAC II computer.

Results have been obtained for several models at both 90 and 296°K. The parameters of these models are listed in Table II, while their agreement with the experimental elastic and dielectric constants is shown in Table III. The agreement with the measured dispersion relations is shown for models I and III in Fig. 7 and for models IV and V in Fig. 9. The dispersion relations in the (1,1,0) and (1,1,1) directions have also been calculated for some of the models and are shown in Figs. 8 and 10.

A particularly interesting feature of the models is that the large charges in the frequencies of the transverse optic mode with lowest frequency have been obtained with comparatively small changes in the parameters of the models. Model VI, for example, has the ionic charges and polarizabilities of model V at 296°K but fits the 90°K measurements with only small changes in the short-range force constants.

TABLE III. The agreement between the experimental elastic and dielectric constants and those calculated using the models.

Model	Elastic constants 10^{12} dynes cm^{-2} C_{11}	C_{12}	C_{44}	Dielectric constants ϵ^0	ϵ^∞
90°K					
Exptl.	3.35[a]	1.05[a]	1.27[a]	1305[b]	5.5[c]
I	2.95	1.00	1.01	193.1	1
II	3.26	1.07	1.80	1161.3	3.88
III	3.19	1.07	1.71	1171.8	3.83
IV	3.40	1.05	1.57	889.9	3.64
V	3.48	1.02	1.33	1227.4	5.66
VI	3.49	1.05	1.33	1150.7	6.30
296°K					
Exptl.	3.30[d]	1.01[d]	1.24[d]	301.0[b]	5.5[c]
I	2.98	1.12	1.02	37.9	1
II	3.30	1.25	1.80	300.9	3.8
III	3.31	1.27	1.71	292.9	3.8
IV	3.45	1.24	1.65	215.1	3.73
V	3.46	1.19	1.28	244.2	5.78

[a] These have been obtained from Eq. (8) neglecting the effects of the 110°K transition.
[b] See Refs. 22 and 23.
[c] See Ref. 25.
[d] See Ref. 11

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

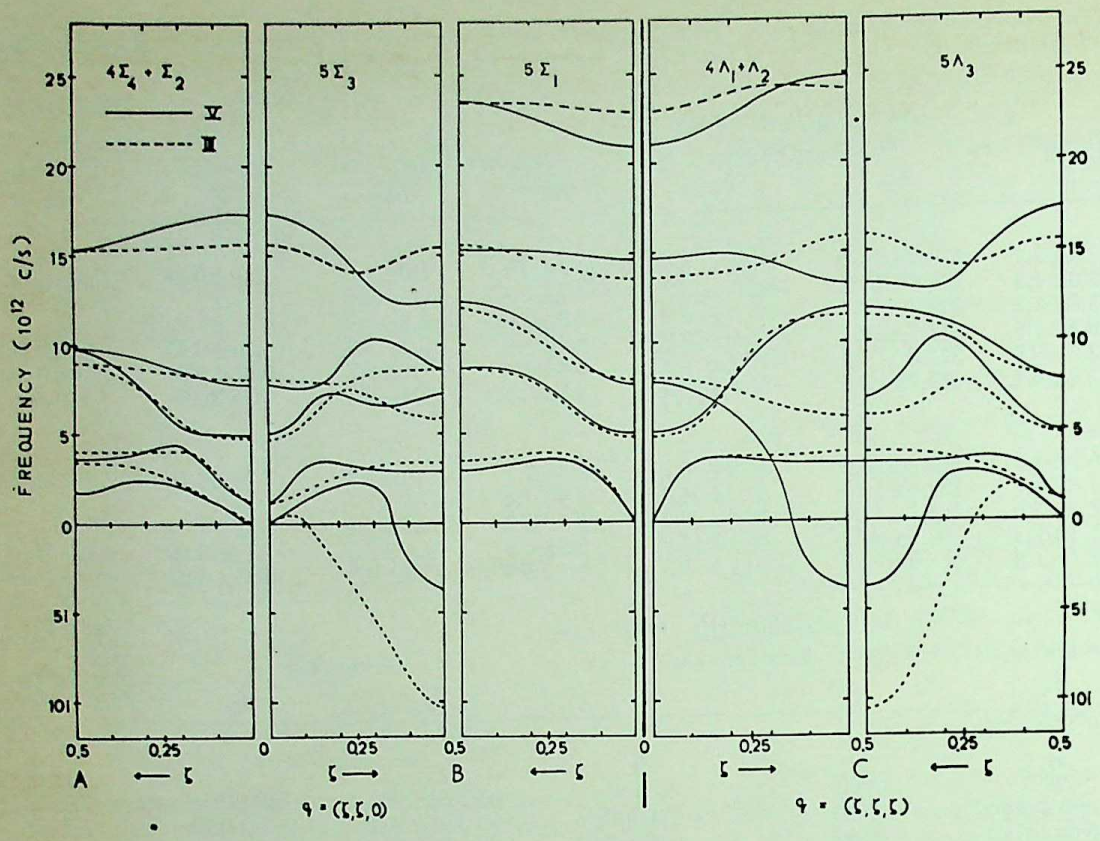


FIG. 8. The dispersion curves in the $(\zeta,\zeta,0)$ and (ζ,ζ,ζ) directions for models III and V at 90°K. The irreducible representations in order of increasing frequency are at the point (A); M_2', M_3', M_5, M_3' (model III), and M_3', M_2', M_5, M_3' (model V). At the point (B) they are M_2, M_5', M_3, M_5', M_1, M_5', M_4 for model III and M_3, M_5', M_2, M_5', M_5', M_1, M_4 for model V, while at the zone boundary (C) they are Γ_{12}', Γ_{15}, Γ_{25}, Γ_{15}, Γ_{25}', Γ_2' (model III), and Γ_{25}, Γ_{15}, Γ_{12}', Γ_{25}', Γ_{15}, Γ_2' (model V).

V. THE 110°K PHASE TRANSITION

1. The Temperature Dependence of the Elastic Constants

The temperature dependence of the elastic constants of crystals has been discussed by several authors.[8,26] The potential energy of the crystal is expanded in a double power series of the deformation parameters and of the phonon coordinates, Eq. (40.4) of Ref. 8.

If this expansion is taken about the equilibrium configuration of the lattice at each temperature, then the isothermal elastic constants can be obtained from[8,26]

$$(\alpha\beta,\gamma\delta)=\Phi_{\alpha\beta\gamma\delta}(--)+\sum_{\mathbf{q}}\sum_{j}\Phi_{\alpha\beta\gamma\delta}(--,\mathbf{q}j-\mathbf{q}j)(\hbar/2(\mathbf{q}j))(2n(\mathbf{q}j)+1)+\sum_{jj'}{}'\sum_{\mathbf{q}}\Phi_{\alpha\beta}(-,\mathbf{q}j-\mathbf{q}j')$$
$$\times\Phi_{\gamma\delta}(-,-\mathbf{q}j\mathbf{q}j')\frac{\hbar}{2\omega(\mathbf{q}j)}\left(\frac{2n(\mathbf{q}j)+1}{\omega^2(\mathbf{q}j)-\omega^2(\mathbf{q}j')}\right)-\sum_{j}\sum_{\mathbf{q}}\Phi_{\alpha\beta}(-,\mathbf{q}j-\mathbf{q}j)\Phi_{\gamma\delta}(-,\mathbf{q}j-\mathbf{q}j)$$
$$\times\left[\frac{\hbar}{4\omega(\mathbf{q}j)^3}(2n(\mathbf{q}j)+1)+\frac{\hbar^2}{2\omega(\mathbf{q}j)^2kT}n(\mathbf{q}j)(n(\mathbf{q}j)+1)\right]. \quad (6)$$

The notation is similar to that of Born and Huang.[8] The elastic constants are then given by

$$C_{\alpha\beta\gamma\delta}{}^{\mathrm{iso}}=(1/4V)[(\alpha\beta,\gamma\delta)+(\alpha\beta,\delta\gamma)+(\beta\alpha,\gamma\delta)+(\beta\alpha,\delta\gamma)].$$

The adiabatic and isothermal elastic constants are related by macroscopic relations, which in Voigt's notation are

$$C_{11}{}^{\mathrm{ad}}-C_{11}{}^{\mathrm{iso}}=C_{12}{}^{\mathrm{ad}}-C_{12}{}^{\mathrm{iso}}$$
$$=(TV\alpha^2/C_V)(C_{11}{}^{\mathrm{iso}}+2C_{12}{}^{\mathrm{iso}})^2,$$
$$C_{44}{}^{\mathrm{ad}}-C_{44}{}^{\mathrm{iso}}=0.$$

The third term in the expression (6) for the elastic constants is particularly of interest. The summation is over all the different pairs of normal modes with a particular wave vector and belonging to different branches of the dispersion relations. Figure 11 shows the longitudinal-acoustic and transverse-optic dispersion curves at 296°K and at 90°K. At the higher temperature the curves are well separated, but at 90°K they appear to be almost degenerate over a considerable region of wave-vector space. Since at zero wave vector the temperature dependence of the transverse optic mode is given by Eq. (5), it might be expected that the temperature dependence of the denominator in the third term of Eq. (6) for these two normal modes is given by

$$\omega^2(\mathbf{q}j)-\omega^2(\mathbf{q}j') \quad \alpha \quad T-T_A,$$

[26] G. Leibfried and W. Ludwig, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1961), Vol. 12.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where T_A is the temperature at which they become degenerate. The above temperature dependence can also be obtained by using a lowest order anharmonic theory of the temperature dependence of the normal modes.

The other factors in the expression for the elastic constants are either independent of temperature or as a first approximation can be written as proportional to temperature in a limited temperature range. The elastic constant is then given by an expression of the type

$$C_{\alpha\beta,\gamma\delta}=A+BT+C/(T-T_A). \quad (7)$$

The coefficients A, B, C can, in principle, be obtained from the matrix elements and summations over the normal modes. The coefficient C would be particularly difficult to evaluate, because it depends on the detailed shape of the dispersion curves near the degeneracy.

The temperature dependence of the elastic constants has been measured above 110°K by Bell and Rupprecht,[11] who obtained

$$C_{11}=3.341\times10^{12}[1-2.62\times10^{-4}(T-T_A)-0.0992/(T-T_A)],$$
$$C_{12}=1.049\times10^{12}[1-1.23\times10^{-4}(T-T_A)+0.1064/(T-T_A)], \quad (8)$$
$$C_{44}=1.267\times10^{12}[1-1.30\times10^{-4}(T-T_A)-0.1242/(T-T_A)],$$

with $T_A=108$°K.

These results are clearly in agreement with the temperature dependence predicted by Eq. (7). Bell and Rupprecht[11] also found that the ultrasonic attenuation increases dramatically near 110°K. The ultrasonic attenuation can be calculated in an exactly similar way to the elastic constant[27] and arises when the frequency Ω of the ultrasonic wave is such that

$$\Omega=\omega(\mathbf{q}j)-\omega(\mathbf{q}j'). \quad (9)$$

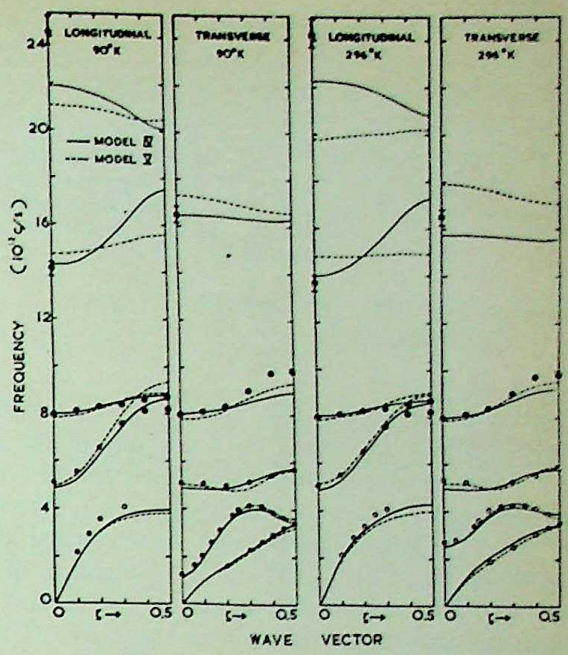


FIG. 9. The dispersion curves for model IV (solid line) and model V (dotted line) for normal modes propagating along the $(\zeta,0,0)$ direction. The experimental measurements are taken from both neutron and infrared spectrometry. (See Refs. 3 and 4.) The irreducible representations at the zone boundaries are for the longitudinal modes M_2', M_1, M_3, M_1, M_2' except for model V at 90°K which gives M_2', M_2, M_1, M_1, M_2' while for the transverse modes they are M_5, M_5', M_5, M_5', M_5. The irreducible representations are listed in order of increasing frequency.

If the transverse optic mode is almost degenerate with the longitudinal acoustic mode in this temperature region over a considerable region of q space, as suggested by Fig. 11, then there will be many normal modes for which Eq. (9) can be satisfied and the ultrasonic attenuation will be large.

2. The Transition

The elastic constants of the crystal will become negative just above T_A, Eqs. (8). When the elastic constants

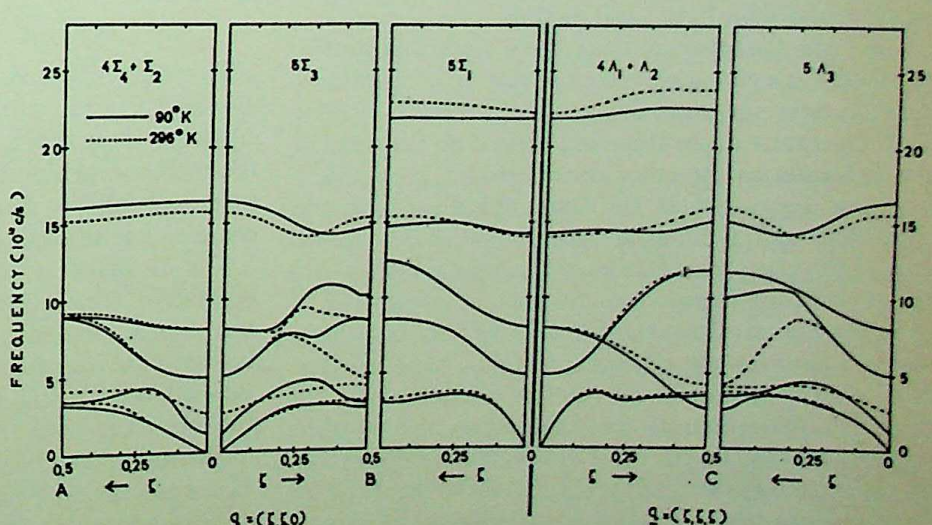


FIG. 10. The dispersion curves of model IV in the $(\zeta,\zeta,0)$ and (ζ,ζ,ζ) directions. In order of increasing frequency the irreducible representations are for the zone boundary (A) M_3', M_2', M_5, M_3' (90°K) and M_2', M_3', M_5, M_3' (296°K), for the zone boundary (B) M_3, M_5', M_5', M_2, M_1, M_5', M_4 (90°K) and M_5', M_3, M_2, M_5', M_1, M_5', M_4, while for the zone boundary (C) they are Γ_{25}, Γ_{15}, Γ_{12}', Γ_{15}, Γ_{25}', Γ_2' (90°K) and Γ_{15}, Γ_{25}, Γ_{12}', Γ_{15}, Γ_{25}', Γ_2' (296°K).

[27] R. A. Cowley, Advan. Phys. 12, 421 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

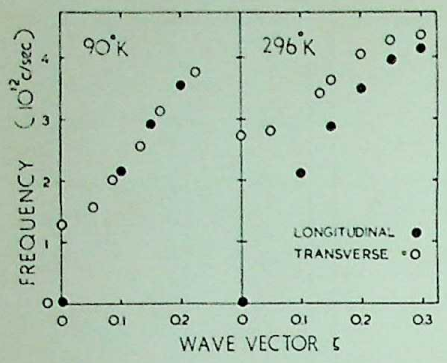


FIG. 11. The longitudinal acoustic and lowest transverse optical branches at 90 and 296°K.

become zero the crystal will undergo a spontaneous deformation to a new structure, the extent of the deformation depending upon the third- and fourth-order elastic constants. Since the elastic constant C_{44} becomes negative at the highest temperature, a tetragonal low-temperature structure can result by combining three shear distortions along different axes in the crystal. A tetragonal structure of strontium titanate below 110°K has been found by Rimai and deMars.[28] In this transition the volume of the crystal does not change, and it is extremely likely that the crystal splits up into small domains, with the major tetrad axis, the z direction in the domain, along different cube axes.

The elastic constants and the free energy of the crystal depend not only on the frequencies $\omega(\mathbf{q}j)$ and $\omega(\mathbf{q}j')$ which are nearly degenerate, but also on the number of these normal modes which are nearly degenerate. The temperature dependence of the distortion can then be estimated if we assume that its magnitude is proportional to the difference in the number of normal modes just above the degenerate frequency to those just below this. The frequencies squared of the transverse optic modes are linearly temperature-dependent, while the deformation is limited by the third-order elastic constants which are assumed to be independent of temperature. The temperature dependence of the distortion is then proportional to $(T_A-T)^{1/2}$ as found experimentally by Rimai and deMars.[28]

The discussion given above is very crude and cannot be expected to give satisfactory results at temperatures far away from the transition temperature. In particular the temperature dependence will depend on the effect of the distortion on the other normal modes, particularly near the degeneracy in the longitudinal acoustic and transverse optical branches, the temperature dependence of the population factors of the normal modes, and the effect of the domain structure on the free energy.

Below the transition the behavior of the elastic constants and ultrasonic attenuation will be very complex. The elastic constants will be an average over the different orientations of the domains, as well as being strongly affected by the exact details of the degeneracies of the normal modes and their behavior in the distorted crystal. The ultrasonic attenuation will also be anomalous because of reflection of the ultrasonic waves at the surfaces of the domains, as suggested by Bell and Rupprecht.[11]

[28] L. Rimai and G. A. deMars, Phys. Rev. 123, 702 (1962).

3. The Dielectric Properties

Below 110°K the frequencies of the normal modes of vibration will be altered by the distortion of the crystal. It is likely that the frequency of the transverse optic mode will be particularly sensitive because of the large changes which occurred when small changes were made in the parameters of the models of Sec. IV. If the lattice spacing decreases by δa along the x and y directions and increases by $2\delta a$ along the unique z axis, the frequency of a transverse optical mode depends upon whether it is polarized along or perpendicular to the z axis.

The changes in the frequencies due to the distortion can be estimated from the volume dependence of the frequencies in the cubic structure. Equation (5) for the frequency will be modified to

$$\omega_z^2=K(T-T_c)+2\delta aB\,,$$
$$\omega_x^2=K(T-T_c)-\delta aB\,,$$

where ω_z is the normal mode polarized along the z axis and $B=d\omega^2/da$.

The distortion of the crystal then changes the Curie temperature by $-2\delta aB/K$ for displacements along the z axis, and by $\delta aB/K$ for those perpendicular to it.

The dielectric constant of a large crystal can then be estimated by assuming that there are an equal number of domains orientated with the z axis along each of the three original cube axes. The average dielectric constant is then

$$\frac{1}{\epsilon^0}=\frac{1}{3}\left(\frac{T-T_c+(2\delta aB/K)}{C}+2\frac{T-T_c-(\delta aB/K)}{C}\right),$$

and therefore,

$$\epsilon^0=C/(T-T_c).$$

The dielectric constant of a bulk sample of the distorted crystal is then identical with that of the crystal if there was no distortion. This result explains the rather surprising experimental result[11,22,23] that the reciprocal of the dielectric constant does not even change slope through the 110°K transition.

The distortion does however influence the low-temperature dielectric properties. Since the Curie temperature is raised in one direction, some of the domains can become polarized above T_c. This spontaneous polarization will occur along different directions in the crystal, corresponding to the different orientations of the domains, and will be frozen in to a particular configuration by the domain pattern produced by the 110°K transition.

The dielectric behavior at low temperatures will therefore be extremely complicated. Deviations from the Curie law will be caused by the spontaneous

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

polarization of some of the domains. Further and probably larger deviations will be caused by quantum effects. Barrett[29] demonstrated how these occur by using a Slater model of a ferroelectric, and very similar conclusions can be obtained by using a lattice dynamical approach. These two effects can probably account for the deviations from a Curie law found experimentally.[22,23]

When an electric field is applied the crystal shows a remanence polarization[22] at temperatures above 50°K, the size of this remanence polarization increasing as the temperature decreases. This polarization may be produced by the electric field aligning the spontaneous polarization of some of the domains, which is then frozen in by the domain structure when the field is released. The increase in magnitude with decreasing temperature arises because more and more domains become spontaneously polarized. The distortion required to shift the Curie temperature to 50°K, an increase of about 15°K, can be estimated from the measurements of the pressure dependence of the Curie temperature,[30] while the sign of the distortion is given by the change in free energy of the crystal. The change in the vibrational part of the free energy due to the changes in the frequency of the transverse optic mode show that the domains are probably elongated along the unique z axis. The dependence of the Curie temperature on lattice spacing is

$$\frac{dT_c}{da}=\frac{1.2\times10^4}{a}.$$

For an increase in Curie temperature of 15°K, the increase in length along the tetragonal axis is 0.005 Å, while the contraction perpendicular to the axis is 0.0025 Å. These distortions are very small, and could easily have been missed in the x-ray diffraction work quoted by Bell and Rupprecht.[11]

VI. DISCUSSION

1. The Parameters of the Models

The parameters for the rigid-ion model (I) show that both the short-range forces between the titanium and oxygen ions and the titanium and oxygen ionic charges are considerably smaller than for the shell models. This decrease in magnitude must in some way compensate for the neglect of the polarizabilities of the ions.

The parameters of all of the shell models are far more similar to one another than they are to the rigid-ion models. The parameters specifying the short-range forces in models II and III, in which the ionic charges were kept fixed, are very similar, but when the ionic charges were allowed to vary in models IV and V there was a decrease in the strontium-oxygen forces and a slight increase in the titanium-oxygen forces. These changes are accompanied by a decrease in the ionic charge of the strontium ions and an increase of the charge on the titanium ions.

It is of interest to compare the parameters of the short-range forces with those obtained by the method suggested by Fowler.[31] Devonshire,[32] Kinase,[33] and Dvorak and Janovec[34] have used these forces to describe the forces in perovskite ferroelectrics. In the notation of Sec. II.2, the parameters of their forces are:

$$A_1=30.1 \quad A_2=429.1 \quad A_3=63.5$$
$$B_1=-2.74 \quad B_2=-41.53 \quad B_3=-6.16.$$

Although these force constants are significantly different from those of any of the models of Sec. IV, they do exhibit some of the principal features; for example, A_2 is by far the largest. Nevertheless, the frequencies of the normal modes of vibration are so sensitive to small changes in the parameters that these short-range force constants do not give an adequate description of the interactions.

When the ionic charges were allowed to vary in the shell models (IV, V), the charge on the titanium ion becomes larger than four, and on the strontium ion about one electronic charge. The charge on the oxygen ions remains at nearly two. These results suggest that the bonding in strontium titanate is more nearly ionic than covalent in character.

There is a considerable reduction in the error of the models when the short-range polarizabilities of the positive ions are allowed to vary (III, IV). These polarizabilities of the positive ions come out either very small or negative, and cannot be readily understood in terms of the shell model. These results are exactly similar to those found for the alkali halides, potassium bromide and sodium iodide.[15] In an exactly analogous way to the alkali halides, when the titanium ion's electrical polarizability was allowed to vary (V) the polarizability of the positive ion was larger than the crystal polarizability,[35] while the polarizability of the oxygen ions was reduced. It was shown[15] that these surprising polarizability parameters could arise in the alkali halides from the neglect of the quadrupole moments produced on the negative ions. A similar explanation can be advanced here in terms of the quadrupole moments of the oxygen ions.

A surprising feature of the polarizabilities of the oxygen ions is their anisotropy. For example, the electrical polarizability of the oxygen ions for the 90°K model V is $2\frac{1}{2}$ times as large in the strontium-oxygen plane as in the titanium-oxygen direction. The short-range polarizability in the titanium-oxygen direction is -1.41, while it is 0.615 in the strontium-oxygen plane.

[29] J. H. Barrett, Phys. Rev. 86, 118 (1952).
[30] W. J. Merz, Phys. Rev. 77, 52 (1950).
[31] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, Cambridge, 1936), 2nd ed.
[32] A. F. Devonshire, Phil. Mag. 40, 1040 (1949).
[33] W. Kinase, Progr. Theoret. Phys. (Kyoto) 15, 529 (1955).
[34] V. Dvorak and V. Janovec, Czech. J. Phys. 812, 461 (1962).
[35] J. R. Tessmann, A. H. Kahn, and W. Shockley, Phys. Rev. 92, 890 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

These polarizability parameters suggest that the electronic wave functions of the oxygen ions may be considerably distorted from spherical, and that the shell parameters Y_3 and k_3, which were assumed to be isotropic, should be different for different directions.

2. The Motion of the Ions in the Normal Modes

There has been considerable discussion as to the nature of the ionic displacements in the transverse optic modes of strontium titanate. The polarization vectors for these normal modes have been obtained for several of the models described earlier, and are listed in Table IV for model IV. The lowest frequency transverse optic mode is found to be predominantly a vibration of the titanium ion against the rest of the structure.

If the crystal becomes unstable against the lowest frequency transverse optic mode of vibration, it might be expected that the displacements of the ions would resemble the pattern of the displacements in the transverse optic mode of vibration.[2] It is therefore of interest to compare the displacements of the ions with the measured displacements of the ions in tetragonal barium titanate,[36] and these are also shown in Table IV. Considering the difference in the masses, sizes and polarizabilities of strontium and barium ions, the displacements of the ions in barium titanate are fairly similar to those in the lowest transverse optic mode of strontium titanate.

3. The Temperature Dependence of the Transverse Optic Modes

One of the most interesting features of strontium titanate is the temperature dependence of the transverse optic mode of lowest frequency. The difference between the parameters of the models which were fitted to the measurements at 90°K and those fitted to the 296°K measurements are not large, and model VI shows that it is possible to obtain the temperature dependence by changing the parameters of the short-range interactions by quite small amounts. These results show that only quite small changes in the parameters of the interactions can give rise to the observed temperature dependence.

The temperature dependence of the normal modes in potassium bromide has been investigated both experimentally[10] and theoretically.[15] The changes in the parameters of the models, which are needed to reproduce the changes in the frequencies of the normal modes from 90 to 400°K, are fractionally very similar to those required for strontium titanate. We can therefore conclude that, broadly speaking, it is not necessary to invoke any mechanism to explain the temperature dependence in strontium titanate, other than what is present in the alkali halides.

The temperature dependence of the normal modes in sodium iodide and potassium bromide has recently been shown to arise largely from the anharmonic interactions between the normal modes.[37] Quantitative calculations of the anharmonic effects in strontium titanate, which will be reported elsewhere, suggest that a similar mechanism is responsible for the temperature dependence of its normal modes.

The anomalous behavior of the normal modes in strontium titanate compared with the alkali halides is associated with the difference in the dynamical matrices of the two crystals. In the alkali halides, the frequencies of the normal modes are all fairly similar in magnitude (apart from the elastic constants region of the acoustic branches). For strontium titanate, on the other hand, there is a very large difference between the frequencies of the optical modes. At 90°K the ratio of the highest to lowest eigenvalue (frequency squared) of the dynamical matrix is as large as 400:1. A change in the parameters of only 0.25% will then hardly alter the largest eigenvalue but can drastically alter the smallest. The ferroelectric character of these perovskites then arises from the almost perfect cancellation of the strong short-range and Coulomb interactions for one particular normal mode. A small change in either of these strong interactions has a drastic effect on the frequencies of this normal mode.

TABLE IV. The displacements of the ions in the $q=0$ modes of model IV at 90°K, polarized along the z direction, and expressed so that the eigenvectors of the dynamical matrix are normalized and orthogonal. The numbering of the modes is in order of increasing frequency.

Model IV	Ti	Sr	O_{III}	O_I	O_{II}
Transverse					
(1)	0.077	0.022	−0.093	−0.129	−0.129
(2)	0.097	−0.074	0.035	0.040	0.040
(3)	0	0	0	−0.177	0.177
(4)	0.007	0.004	−0.217	0.087	0.087
Longitudinal					
(1)	0.080	−0.077	0.050	0.065	0.065
(2)	0	0	0	−0.177	0.177
(3)	0.057	0.001	0.107	−0.143	−0.143
(4)	0.075	0.008	−0.207	−0.032	−0.032
Displacements of ions in $BaTiO_3$ transition[a] (Å)	0.053	0.003	−0.051	−0.047	−0.047

[a] See Ref. 36.

4. Unstable Normal Modes in the (1,1,0) and (1,1,1) Directions

Several of the models described in Sec. IV show instabilities against some of the normal modes of vibration propagating along the (1,1,0) and (1,1,1) directions. Examples of these are shown in Fig. 8 for models III and V. It is unlikely that these instabilities occur in strontium titanate and in this respect those models

[36] B. C. Frazer, H. R. Danner, and R. Pepinsky, Phys. Rev. **100**, 745 (1955).

[37] R. A. Cowley, Phys. Chem. Solids (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which show them are unrealistic. However, these instabilities may well occur in some of the perovskite materials which have phases that are slightly distorted forms of the cubic perovskite structure. For example, $CdTiO_3$, $NaTaO_3$, $CaTiO_3$, $NaNbO_3$ show multiple unit cells which are $2\times2\times2$ times the cubic cell when referred to the original cube axes while $NaNbO_3$ shows a $4\times2\times2$ unit cell, and $PbZrO_3$ shows a $4\times4\times2$ unit cell.[38]

All of the unstable normal modes are rotations of groups of the oxygen ions. In models I and V octahedra rotate around the titanium ion while in model III squares of oxygens rotate about the points $(\frac{1}{2},0,0)$, $(0,\frac{1}{2},0)$ and $(0,0,\frac{1}{2})$. A further feature of these unstable normal modes is that their frequencies are very dependent upon the details of the models. For example, the normal mode Γ_{25}' are unstable for models I and V but for model III have a frequency of 2.6 and for model IV of 3.5.

The transitions to the multiple unit cell perovskites may well be instabilities against these normal modes as suggested by Cochran.[2] Unfortunately the displacements of the ions in real distorted perovskite structures are far more complex than mere oxygen octahedra rotations. However, in several of these materials, the rotations of the oxygen octahedra are the predominant displacements.[39] The displacements of the other ions are not, however, negligible and a detailed treatment would be very complicated. The unstable normal mode must be strongly coupled both to the macroscopic strain parameters and to some of the other normal modes in the crystal.

5. The Models of Other Authors

Several authors have previously used models for the lattice dynamics of perovskite ferroelectrics. Devonshire[32] used a rigid-ion model with Fowler's[31] short-range forces to attempt to relate the parameters of his thermodynamic theory to the microscopic interactions. Recently Dvorak and Janovec[34] have calculated the frequencies of the transverse optic modes of this model, and have found the structure unstable unless the formal ionic charges are multiplied by a factor 0.16. The frequencies are not then in agreement with the results of neutron or infrared spectrometry.

Kinase[33] has used this model, and included the electrical polarizabilities of the ions, but the results were no more satisfactory.

Last[18] suggested that the infrared spectra could be explained by assuming that the titanium-oxygen octahedra was a strongly bound complex which was weakly coupled to the other anion. As the lowest frequency optic mode is a vibration of the titanium against the rest of the structure this model is unsatisfactory. Rajopal and Srinivasan[40] studied the frequencies of vibration at $\mathbf{q}=0$ when all the Coulomb forces are neglected. The high- and low-frequency dielectric constants of this model are then equal (to unity!) and the frequencies do not agree with the experimental results. A similar model has been used by Silverman and Joseph[41] for the transverse optic modes with $\mathbf{q}=0$.

VII. CONCLUSIONS

The frequency versus wave-vector dispersion curves of strontium titanate for several normal modes propagating in the [0,0,1] direction have been measured experimentally by using neutron spectrometry. These results have been used to find the parameters of several models of the crystal within the harmonic approximation. It was found that quite reasonable agreement with experiment could be obtained by using shell models in which the ions interact with one another both through short-range forces between neighboring ions, and through long-range electrostatic forces, determined by the ionic charges. The effects of the polarizabilities of the ions are also included. The principal features of the more successful models are that the titanium-oxygen short-range forces are very large, while the ionic charges on the ions are also large and very nearly the formal charges.

Of particular interest in strontium titanate is the temperature dependence of the transverse optic mode with lowest frequency. This was measured experimentally and for the $\mathbf{q}=0$ mode the square of the frequency was found to be proportional to temperature above 90°K, in agreement with the temperature dependence of the static dielectric constant as predicted by Cochran.[2] This result shows that it is a valid approach to treat the problem of ferroelectricity as an instability of the crystal against one of the normal modes. The motion of the ions in this normal mode was obtained for several of the models used, and was found to be predominantly a vibration of the titanium ion against the oxygen octahedron. The displacements of the ions at the ferroelectric transition in barium titanate are very similar to this.

It was found that the experimentally observed changes in the frequencies of this normal mode could be obtained by changing the parameters of the shell models only slightly. The fractional changes were very similar to those needed to reproduce the temperature dependence of the normal modes in the alkali halides. This result suggests that it is unnecessary to introduce any mechanism to explain the temperature dependence of this normal mode, other than those which occur in the alkali halides. The indications are then that the temperature dependence of this normal mode can be explained

[38] F. Jona and G. Shirane, *Ferroelectric Crystals* (Pergamon Press, Inc., New York, 1963).
[39] H. D. Megaw, *Ferroelectricity in Crystals* (Methuen and Company, Ltd., London, 1957).
[40] R. K. Rajopal and R. Srinivasan, Phys. Chem. Solids **23**, 633 (1962).
[41] B. D. Silverman and R. I. Joseph, Technical Memorandum T-451, Raytheon Company, 1963 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in terms of the anharmonic interactions between the normal modes of vibration.

The dispersion curves along the (1,1,0) and (1,1,1) directions of the models have been calculated and some of these show normal modes against which the crystal would be unstable. Although it is unlikely that these instabilities occur in strontium titanate, they may well be associated with the transitions in other perovskite materials.

The transition in strontium titanate at 110°K has been discussed in terms of an accidental degeneracy of two branches of the dispersion curves. The experimental results from neutron spectrometry show that this degeneracy occurs very near 110°K. The anomalous elastic properties near 110°K and the smooth behavior of the static dielectric constant can be easily obtained with this theory. It is suggested that the conflicting experimental results on the low-temperature dielectric properties, and the temperature dependence of the distortion of the crystal can also be explained by this approach.

It is clearly of interest to try to obtain more satisfactory models of strontium titanate, in the hope of obtaining a better understanding of both the temperature dependence of the static dielectric constant and of the 110°K transition. Experimental measurements of the dispersion curves in the (1,1,0) and (1,1,1) directions would enable the parameters of the models to be fitted more accurately. The models could then be extended by including anisotropic oxygen polarizabilities, more short-range forces, and abandoning the assumption that $\mathbf{R}=\mathbf{T}=\mathbf{S}$.

The difficulties in describing the results in terms of a harmonic model might then become more serious, and it might well prove necessary to use a complete anharmonic theory to describe the lattice vibrations.

ACKNOWLEDGMENTS

The author is grateful for helpful discussions on the experimental measurements with Professor B. N. Brockhouse, on group theory with Dr. V. Heine, and on distorted perovskite structures with Dr. H. D. Megaw and Dr. I. Lefkowitz. Above all, however, he is indebted to Dr. W. Cochran, F.R.S., who suggested the study of strontium titanate and who has continually given helpful advice and criticism.

The author is grateful for the award of a NATO studentship from the Department of Scientific and Industrial Research and to Atomic Energy of Canada Ltd. for allowing him to use experimental facilities. Permission to use the EDSAC II computer at the Cambridge University Mathematical Laboratory was given by Dr. M. V. Wilkes, F.R.S., for which the author is grateful. The excellent single crystal of strontium titanate was provided by the Titanium Division of the National Lead Company, for which the author is grateful.

APPENDIX I

A. The Characters of the Symmetry Operations

The representation chosen for the normal modes is obtained by displacing each ion in turn along the coordinate axes by a unit distance. The characters of the different symmetry elements are:

identity operator

$$\chi(E)=15,$$

triad axis

$$\chi(C_3)=0,$$

tetrad axis along the z axis

$$\chi(C_{4z})=1+2\exp(i2q_xr),$$

diad axis along the z axis

$$\chi(C_{2z})=-1-\exp(i2q_xr)-\exp(i2q_yr)-2\exp(i2(q_x+q_y)r),$$

diad axis at 45° to both x and y axes

$$\chi(C_{2xy}')=-2-\exp(i2q_zr),$$

inversion operator

$$\chi(I)=-3[1+\exp(i2(q_x+q_y+q_z)r)+\exp(i2(q_x+q_y)r)+\exp(i2(q_y+q_z)r)+\exp(i2(q_x+q_z)r)],$$

plane of symmetry perpendicular to the z axis

$$\chi(\sigma_{xy})=2+3\exp(i2q_zr),$$

plane of symmetry at 45° to x and y axes and including z axis

$$\chi(\sigma_d)=3,$$

inversion tetrad axis along the z axis

$$\chi(S_{4z})=-1-\exp(i2q_zr)-\exp(i2(q_x+q_z)r),$$

inversion hexad axis

$$\chi(S_6)=0.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B. The Displacements of the Ions in Different Normal Modes

Irreducible representation	Normal mode
$\zeta=(0,0,\zeta)$	
Δ_1	Sr_z, Ti_z, O_{Iz}, $O_{IIz}=O_{IIIz}$.
Δ_2	$O_{IIz}=-O_{IIIz}$.
Δ_5	Sr_x, Ti_xO_{Ix}, O_{IIx}, O_{IIIx}. Sr_y, Ti_y, O_{Iy}, O_{IIy}, O_{IIIy}.
$\zeta=(0,0,\frac{1}{2})$	
M_1	Ti_z, $O_{IIz}=O_{IIIz}$.
M_2'	Sr_z, O_{Iz}.
M_3	$O_{IIz}=-O_{IIIz}$.
M_5	Ti_x, O_{IIx}, O_{IIIx}. Ti_y, O_{IIy}, O_{IIIy}.
M_5'	Sr_x, O_{Ix}. Sr_y, O_{Iy}.
$\zeta=(\zeta,\zeta,0)$	
Σ_1	$Ti_x=Ti_y$, $Sr_x=Sr_y$, $O_{Ix}=O_{Iy}$, $O_{IIx}=O_{IIIy}$, $O_{IIIx}=O_{IIy}$.
Σ_2	$O_{IIz}=-O_{IIIz}$.
Σ_3	$Ti_x=-Ti_y$, $Sr_x=-Sr_y$, $O_{Ix}=-O_{Iy}$, $O_{IIx}=-O_{IIIy}$, $O_{IIIx}=-O_{IIy}$.
Σ_4	Ti_z, Sr_z, O_{Iz}, $O_{IIz}=O_{IIIz}$.
$\zeta=(\frac{1}{2},\frac{1}{2},0)$	
M_1	$O_{IIx}=O_{IIIy}$.
M_2	$O_{IIIx}=-O_{IIy}$.
M_3	$O_{IIx}=-O_{IIIy}$.
M_4	$O_{IIIx}=O_{IIy}$.
M_2'	Sr_z.
M_3'	Ti_z, O_{Iz}.
M_5	O_{IIz}. O_{IIIz}.
M_5'	Ti_x, Sr_y, O_{Ix}. Ti_y, Sr_x, O_{Iy}.
$\zeta=(\zeta,\zeta,\zeta)$	
Λ_1	$Sr_x=Sr_y=Sr_z$, $Ti_x=Ti_y=Ti_z$, $O_{IIIx}=O_{IIy}=O_{Iz}$, $O_{IIIy}=O_{IIIz}$ $=O_{IIz}=O_{IIz}=O_{Ix}=O_{Iy}$.
Λ_2	$O_{Iz}=O_{IIz}=O_{IIIy}=-O_{Iy}$ $=-O_{IIx}=-O_{IIIz}$.
Λ_3	$Ti_x=-(Ti_y+Ti_z)$, $Sr_x=-(Sr_y+Sr_z)$, $O_{IIIx}=-(O_{IIy}+O_{Iz})$, $O_{Ix}=-(O_{IIIy}+O_{IIz})$,

Irreducible representation	Normal mode
$\zeta=(\zeta,\zeta,\zeta)$	
Λ_3	$O_{IIx}=-(O_{Iy}+O_{IIIz})$. $Ti_x=-Ti_y$, $Sr_z=-Sr_y$, $O_{Iz}=-O_{IIy}$, O_{IIz} $=-O_{IIIy}$, $O_{IIIz}=-O_{Iy}$.
$\zeta=(\frac{1}{2},\frac{1}{2},\frac{1}{2})$	
Γ_2'	$O_{Iz}=O_{IIy}=O_{IIIx}$.
Γ_{12}'	$O_{Iz}=O_{IIy}=-\frac{1}{2}O_{IIIx}$. $O_{Iz}=-O_{IIy}$.
Γ_{25}	$O_{Iy}=-O_{IIz}$. $O_{Ix}=-O_{IIIz}$. $O_{IIx}=-O_{IIIy}$.
Γ_{25}'	Ti_z. Ti_y. Ti_z.
Γ_{15}	Sr_x, $O_{Iy}=O_{IIz}$. Sr_y, $O_{Iz}=O_{IIIz}$. Sr_z, $O_{IIx}=O_{IIIy}$.
$\zeta=(\frac{1}{2},\frac{1}{2},\zeta)$	
Δ_1	Sr_z, $O_{IIx}=O_{IIIy}$.
Δ_1'	$O_{IIIx}=-O_{IIy}$.
Δ_2	$O_{IIx}=-O_{IIIy}$.
Δ_2'	Ti_z, $O_{IIIx}=O_{IIy}$, O_{Iz}.
Δ_5	Ti_x, Sr_y, O_{Ix}, O_{IIIz}. Ti_y, Sr_x, O_{Iy}, O_{IIz}.
$\zeta=(\zeta,\zeta,\frac{1}{2})$	
Σ_1	Ti_z, $Sr_x=Sr_y$, $O_{IIIz}=O_{IIz}$, $O_{Iy}=O_{Iz}$.
Σ_3	$Sr_x=-Sr_y$, $O_{IIIz}=-O_{IIz}$, $O_{Iy}=-O_{Ix}$.
Σ_4	$Ti_x=Ti_y$, Sr_z, O_{Iz}, $O_{IIy}=O_{IIIx}$, $O_{IIz}=O_{IIIy}$.
Σ_2	$Ti_x=-Ti_y$, $O_{IIIx}=-O_{IIy}$, $O_{IIx}=-O_{IIIy}$.
$\zeta=(\frac{1}{2},0,\zeta)$	
Σ_1	Ti_z, Sr_z, O_{Ix}, O_{IIx}, O_{IIIz}.
Σ_4	Sr_y, O_{IIIy}.
Σ_3	Ti_z, Sr_x, O_{IIIx}, O_{IIz}, O_{Iz}.
Σ_2	Ti_y, O_{Iy}, O_{IIy}.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

C. The Compatibility Relations

$\zeta=(0,0,\zeta)$ with $\zeta=(0,0,0)$:

$\Gamma_{15}-\Delta_2+\Delta_5$, $\Gamma_{25}-\Delta_2+\Delta_5$.

$\zeta=(0,0,\zeta)$ with $\zeta=(0,0,\frac{1}{2})$:

Δ_1-M_1, Δ_1-M_2', Δ_2-M_3, Δ_5-M_5, Δ_5-M_5'.

$\zeta=(\zeta,\zeta,0)$ with $\zeta=(0,0,0)$:

$\Gamma_{15}-\Sigma_1+\Sigma_3+\Sigma_4$, $\Gamma_{25}-\Sigma_1+\Sigma_3+\Sigma_2$.

$\zeta=(\zeta,\zeta,0)$ with $\zeta=(\frac{1}{2},\frac{1}{2},0)$:

$M_1-\Sigma_1$, $M_2-\Sigma_3$, $M_2'-\Sigma_4$, $M_3-\Sigma_3$, $M_3'-\Sigma_4$, $M_4-\Sigma_1$, $M_5-\Sigma_2+\Sigma_4$, $M_5'-\Sigma_1+\Sigma_3$.

$\zeta=(\zeta,\zeta,\zeta)$ with $\zeta=(0,0,0)$:

$\Gamma_{15}-\Lambda_1+\Lambda_3$, $\Gamma_{25}-\Lambda_2+\Lambda_3$.

$\zeta=(\zeta,\zeta,\zeta)$ with $\zeta=(\frac{1}{2},\frac{1}{2},\frac{1}{2})$:

$\Gamma_{25}'-\Lambda_1+\Lambda_3$, $\Gamma_{15}-\Lambda_1+\Lambda_3$, $\Gamma_{25}-\Lambda_2+\Lambda_3$, $\Gamma_2'-\Lambda_1$, $\Gamma_{12}'-\Lambda_3$.

$\zeta=(\frac{1}{2},\frac{1}{2},\zeta)$ with $\zeta=(\frac{1}{2},\frac{1}{2},0)$:

$M_1-\Delta_1$, $M_2-\Delta_1'$, $M_2'-\Delta_1$, $M_3-\Delta_2$, $M_3'-\Delta_2'$, $M_4-\Delta_2'$, $M_5-\Delta_5$, $M_5'-\Delta_5$.

$\zeta=(\frac{1}{2},\frac{1}{2},\zeta)$ with $\zeta=(\frac{1}{2},\frac{1}{2},\frac{1}{2})$:

$\Gamma_2'-\Delta_2'$, $\Gamma_{12}'-\Delta_1'+\Delta_2'$, $\Gamma_{25}-\Delta_2+\Delta_5$, $\Gamma_{15}-\Delta_1+\Delta_5$, $\Gamma_{25}'-\Delta_2'+\Delta_5$.

$\zeta=(\zeta,\zeta,\frac{1}{2})$ with $\zeta=(0,0,\frac{1}{2})$:

$M_1-\Sigma_1$, $M_2'-\Sigma_4$, $M_3-\Sigma_3$, $M_5-\Sigma_2+\Sigma_4$, $M_5'-\Sigma_1+\Sigma_3$.

$\zeta=(\zeta,\zeta,\frac{1}{2})$ with $\zeta=(\frac{1}{2},\frac{1}{2},\frac{1}{2})$:

$\Gamma_2'-\Sigma_4$, $\Gamma_{12}'-\Sigma_2+\Sigma_4$, $\Gamma_{25}-\Sigma_1+\Sigma_2+\Sigma_3$, $\Gamma_{15}-\Sigma_1+\Sigma_3+\Sigma_4$, $\Gamma_{25}'-\Sigma_1+\Sigma_2+\Sigma_4$.

$\zeta=(\frac{1}{2},0,\zeta)$ with $\zeta=(\frac{1}{2},0,0)$:

$M_1-\Sigma_1$, $M_2'-\Sigma_3$, $M_3-\Sigma_1$, $M_5-\Sigma_2+\Sigma_3$, $M_5'-\Sigma_4+\Sigma_1$.

$\zeta=(\frac{1}{2},0,\zeta)$ with $\zeta=(\frac{1}{2},0,\frac{1}{2})$:

$M_1-\Sigma_1$, $M_2-\Sigma_3$, $M_2'-\Sigma_4$, $M_3-\Sigma_1$, $M_3'-\Sigma_2$, $M_4-\Sigma_3$, $M_5-\Sigma_2+\Sigma_4$, $M_5'-\Sigma_1+\Sigma_3$.

APPENDIX II: THE CONTRIBUTION OF THE SHORT-RANGE FORCES TO THE MATRIX $R(KK')$

The complete (15×15) matrix can be split up into the (5×5) matrices for the (xx), $(yy)\cdots$ components. In this appendix the (xx) and (xy) components are tabulated, in terms of the force constants introduced in Sec. II.2. C_x, C_y, C_z are the cosines of q_xr, q_yr, q_zr, while S_x, S_y, S_z are the analogous sines, and $2r$ is the lattice parameter.

The matrix $\mathbf{R}_{xx}$ in units of e^2/v is

	Ti	Sr	O_I	O_{II}	O_{III}
Ti	A_2+2B_2	0	$-B_2C_z$	$-B_2C_y$	$-A_2C_x$
Sr	0	$2A_1+4B_1$	$-(A_1+B_1)C_xC_y$	$-(A_1+B_1)C_xC_z$	$-2B_1C_yC_z$
O_I	$-B_2C_z$	$-(A_1+B_1)C_xC_y$	$A_1+B_1+B_2+A_3+3B_3$	$-2B_3C_yC_z$	$-(A_3+B_3)C_xC_z$
O_{II}	$-B_2C_y$	$-(A_1+B_1)C_xC_z$	$-2B_3C_yC_z$	$A_1+B_1+B_2+A_3+3B_3$	$-(A_3+B_3)C_xC_y$
O_{III}	$-A_2C_x$	$-2B_1C_yC_z$	$-(A_3+B_3)C_xC_z$	$-(A_3+B_3)C_xC_y$	$2B_1+A_2+2A_3+2B_3$

The matrix $\mathbf{R}_{xy}$ is, in units e^2/v:

	Ti	Sr	O_I	O_{II}	O_{III}
Ti	0	0	0	0	0
Sr	0	0	$(A_1-B_1)S_xS_y$	0	0
O_I	0	$(A_1-B_1)S_xS_y$	0	0	0
O_{II}	0	0	0	0	$(A_3-B_3)S_xS_y$
O_{III}	0	0	0	$(A_3-B_3)S_xS_y$	0

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX III

TABLE V. Coulomb coefficients.[a]

$\zeta = (0,0,\zeta)$

Ti-Ti

ζ	xx	yy	zz
0.5	−4.844	−4.844	9.687
0.4	−4.782	−4.782	9.563
0.3	−4.618	−4.618	9.236
0.2	−4.416	−4.416	8.832
0.1	−4.252	−4.252	8.503
0	−4.189	−4.189	8.378

Ti-Sr

ζ	xx	yy	zz
0.5	0	0	0
0.4	−1.281	−1.281	2.561
0.3	−2.443	−2.443	4.886
0.2	−3.375	−3.375	6.750
0.1	−3.979	−3.979	7.958
0	−4.189	−4.189	8.378

Ti-O_I

ζ	xx	yy	zz
0.5	0	0	0
0.4	3.339	3.339	−6.678
0.3	6.359	6.359	−12.717
0.2	8.765	8.765	−17.530
0.1	10.316	10.316	−20.632
0	10.852	10.852	−21.704

Ti-O_{III}

ζ	xx	yy	zz
0.5	−33.622	10.323	23.299
0.4	−33.684	10.373	23.311
0.3	−33.845	10.505	23.341
0.2	−34.046	10.668	23.377
0.1	−34.208	10.801	23.407
0	−34.271	10.852	23.419

Sr-O_I

ζ	xx	yy	zz
0.5	−7.996	−7.996	15.992
0.4	−8.046	−8.046	16.091
0.3	−8.177	−8.177	16.354
0.2	−8.340	−8.340	16.679
0.1	−8.472	−8.472	16.944
0	−8.523	−8.523	17.045

Sr-O_{III}

ζ	xx	yy	zz
0.5	0	0	0
0.4	1.370	−2.619	1.249
0.3	2.614	−4.989	2.376
0.2	3.610	−6.880	3.271
0.1	4.255	−8.101	3.846
0	4.479	−8.523	4.044

$\zeta = (\zeta,\zeta,0)$

Ti-Ti

ζ	xx	yy	zz	xy	yz	xz
0.5	2.677	2.677	−5.354	0	0	0
0.4	2.627	2.627	−5.255	0.716	0	0
0.3	2.492	2.492	−4.985	2.466	0	0
0.2	2.313	2.313	−4.626	4.394	0	0
0.1	2.157	2.157	−4.313	5.787	0	0
0	2.094	2.094	−4.189	6.283	0	0

Ti-Sr

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	10.620	0	0
0.4	0.196	0.196	−0.392	10.094	0	0
0.3	0.713	0.713	−1.426	8.841	0	0
0.2	1.360	1.360	−2.720	7.515	0	0
0.1	1.890	1.890	−3.780	6.599	0	0
0	2.094	2.094	−4.189	6.283	0	0

Ti-O_I

ζ	xx	yy	zz	xy	yz	xz
0.5	16.559	16.559	−33.118	0	0	0
0.4	16.608	16.608	−33.215	0.702	0	0
0.3	16.741	16.741	−33.482	2.426	0	0
0.2	16.919	16.919	−33.838	4.353	0	0
0.1	17.074	17.074	−34.147	5.771	0	0
0	17.135	17.135	−34.271	6.283	0	0

Ti-O_{III}

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	0	0	0
0.4	−9.436	6.240	3.196	2.003	0	0
0.3	−17.483	11.323	6.160	3.779	0	0
0.2	−23.351	14.732	8.618	5.147	0	0
0.1	−26.838	16.570	10.268	5.997	0	0
0	−27.988	17.135	10.852	6.283	0	0

Sr-O_I

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	17.060	0	0
0.4	−0.210	−0.210	0.420	15.924	0	0
0.3	−0.763	−0.763	1.526	13.071	0	0
0.2	−1.455	−1.455	2.909	9.756	0	0
0.1	−2.021	−2.021	4.043	7.221	0	0
0	−2.239	−2.239	4.479	6.283	0	0

Sr-O_{III}

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	0	0	0
0.4	4.013	−1.535	−2.477	1.852	0	0
0.3	7.221	−2.429	−4.792	3.568	0	0
0.2	9.316	−2.581	−6.734	4.991	0	0
0.1	10.423	−2.370	−8.053	5.946	0	0
0	10.762	−2.240	−8.523	6.283	0	0

$\zeta = (\zeta,\zeta,\zeta)$

Ti-Ti

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	0	0	0
0.4	0	0	0	0.283	0.283	0.283
0.3	0	0	0	1.170	1.170	1.170
0.2	0	0	0	2.499	2.499	2.499
0.1	0	0	0	3.706	3.706	3.706
0	0	0	0	4.189	4.189	4.189

Ti-Sr

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	0	0	0
0.4	0	0	0	2.923	2.923	2.923
0.3	0	0	0	4.655	4.655	4.655
0.2	0	0	0	4.934	4.934	4.934
0.1	0	0	0	4.472	4.472	4.472
0	0	0	0	4.189	4.189	4.189

Ti-O_{III}

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	0	0	0
0.4	−9.860	4.930	4.930	1.121	0.151	1.121
0.3	−18.557	9.278	9.278	2.198	0.987	2.198
0.2	−25.047	12.524	12.524	3.178	2.404	3.178
0.1	−28.856	14.428	14.428	3.911	3.686	3.911
0	−30.082	15.041	15.041	4.189	4.189	4.189

Sr-O_I

ζ	xx	yy	zz	xy	yz	xz
0.5	0	0	0	14.461	0	0
0.4	−0.553	−0.553	1.107	13.462	0.466	0.466
0.3	−1.858	−1.858	3.715	10.842	1.618	1.618
0.2	−3.189	−3.189	6.379	7.641	2.908	2.908
0.1	−4.052	−4.052	8.104	5.125	3.851	3.851
0	−4.334	−4.334	8.668	4.189	4.189	4.189

[a] The dimensionless Coulomb coefficients are defined in Ref. 16. If these numbers are multiplied by $(e^2/v)Z_KZ_{K'}$ they give the Coulomb contributions to the dynamical matrix.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Optical Absorption in an Electric Field

JOSEPH CALLAWAY*
Aeronutronic Division of Philco Corporation, Newport Beach, California
(Received 26 December 1963)

A previous calculation of the effect of an external, uniform electric field on the optical absorption associated with a direct transition between bands is extended. The absorption is shown to consist of a series of steps whose width is determined by the separation of the discrete levels produced by the field.

IN a previous calculation[1] (which will be referred to as I), it was shown that the discrete levels which are formed when a uniform electric field is present in an insulating crystal[2] produce a periodic structure in the energy dependence of the absorption due to interband transitions. The expression for the absorption coefficient which was given in that paper is a slowly convergent Fourier series which cannot easily be interpreted. It is the object of the present work to transform this series in such a way that the energy dependence of the optical absorption is evident. I hope that this work will help to stimulate experimental investigation of the discrete levels.

An infinite crystal containing a uniform electric field $\mathcal{E}$, in the x direction is considered. From Eqs. (14) and (21) of I, the following expression is obtained for the absorption coefficient α, produced by a direct transition between bands n' and n (the notation conforms to I as far as possible):

$$\alpha=\frac{2\kappa^2}{F\pi\omega n\epsilon_0 c\mathfrak{A}_0^2}\left[\int\frac{d\mathbf{k}_\perp}{(2\pi)^2}|M_{nn'}(x_0)|^2+2\sum_{l=1}^{\infty}\int\frac{d\mathbf{k}_\perp}{(2\pi)^2}|M_{nn'}(x_0)|^2\cos 2\pi l x_0\right], \quad (1)$$

in which κ is the width of the Brillouin zone in the x direction (which coincides with a reciprocal lattice vector); $\mathbf{k}_\perp$ represents components of the wave vector perpendicular to the field; $\mathfrak{A}_0$ is the magnitude of the vector potential of the radiation field (which has circular frequency ω); n is the index of refraction, and ϵ_0 is the permittivity of free space. In addition, $F=e\mathcal{E}$, and

$$x_0=\frac{\kappa}{2\pi F}(\Delta_{nn'}-\hbar\omega), \quad (2a)$$

where

$$\Delta_{nn'}(\mathbf{k}_\perp)=\kappa^{-1}\int_{-\kappa/2}^{\kappa/2}[E_n(k_x,\mathbf{k}_\perp)-E_{n'}(k_x,\mathbf{k}_\perp)]dk_x. \quad (2b)$$

The matrix element $M_{nn'}$ is given by

$$M_{nn'}=\frac{2\pi e\mathfrak{A}_0\beta^{1/3}}{m\kappa}(\boldsymbol{\varepsilon}\cdot\mathbf{p}_{nn'})\,\mathrm{Ai}(\sigma\beta^{1/3}), \quad (3)$$

where $\mathbf{p}_{nn'}$ is the usual interband matrix element of the momentum and $\boldsymbol{\varepsilon}$ is the polarization vector of the radiation. Both bands are treated in the effective mass approximation, so that the energy difference between bands is

$$E_n-E_{n'}=E_g+(\hbar^2k^2/2\mu). \quad (4)$$

In this equation, E_g is the energy gap at $k=0$, and μ is the reduced mass for the two bands: $\mu^{-1}=m_n^{-1}+m_{n'}^{-1}$. We also have

$$\beta=2\mu F/\hbar^2;\quad \sigma=F^{-1}(E_g-\hbar\omega+\hbar^2k_\perp^2/2\mu),$$

and Ai is the Airy integral, which is defined by

$$\mathrm{Ai}(z)=\frac{1}{\pi}\int_0^\infty\cos(sz+s^3/3)ds.$$

Other quantities which are not explicitly defined above have their usual significance.

In I, approximate expressions were worked out for the terms of the series, Eq. (1). It is, however, evident from Eqs. (1) and (2) without approximation that whenever the photon energy $\hbar\omega$ increases by $\Delta(\hbar\omega)$ with

$$\Delta(\hbar\omega)=2\pi F/\kappa, \quad (5)$$

all the cosines return to their original values. This energy change is just the separation between the discrete levels so that the absorption must contain some structure with this period. In order to investigate this structure, we proceed as follows. The sum of trigonometric functions is evaluated using the identity[3]

$$1+2\sum_{l=1}^{\infty}\cos ly=2\pi\sum_{m=-\infty}^{\infty}\delta(y-2m\pi). \quad (6)$$

Then with the use of (3)

$$\alpha=K\frac{2\mu}{\omega\beta^{1/3}}\int d\mathbf{k}_\perp\,\mathrm{Ai}^2(\sigma\beta^{1/3})\left[1+2\sum_{l=1}^{\infty}\cos 2\pi l x_0\right]$$

$$=(4\pi K\mu/\beta^{1/3}\omega)\sum_{j=-\infty}^{\infty}\int d\mathbf{k}_\perp\,\mathrm{Ai}^2(\sigma\beta^{1/3})\delta(2\pi x_0-2\pi j), \quad (7)$$

---

* Consultant. Permanent address: Department of Physics, University of California, Riverside, California.

[1] J. Callaway, Phys. Rev. 130, 549 (1963). Other calculations of the effect of an electric field on optical absorption have been reported by: W. Franz, Z. Naturforsch. 13, 484 (1958); L. V. Keldysh, Zh. Eksperim. i Teor. Fiz. 34, 1138 (1958) [English transl.: Soviet Phys.—JETP 7, 788 (1958)]; and K. Tharmalingam, Phys. Rev. 130, 2204 (1963).

[2] G. Wannier, Phys. Rev. 117, 432 (1960).

[3] G. Goertzel and N. Tralli, *Some Mathematical Methods of Physics* (McGraw-Hill Book Company, Inc., New York, 1960), p. 123.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in which

$$K=2e^2|\boldsymbol{\varepsilon}\cdot\mathbf{p}_{nn'}|^2/\pi\hbar^2m^2n\epsilon_0c. \tag{8}$$

From Eqs. (2) and (4), we obtain for x_0

$$2\pi x_0=\kappa(\sigma+\chi\beta^{-1/3}), \tag{9}$$

in which

$$\chi=\kappa^2/12\beta^{2/3}.$$

Then we change the variable of integration in (7) to $y=\sigma\beta^{1/3}$ and obtain

$$\alpha=\frac{4\pi^2K\mu\beta^{2/3}}{\omega\kappa}\sum_{j=-\infty}^{\infty}\int dy\,\mathrm{Ai}^2(y)$$

$$\times\delta[y+\chi-(2\pi j\beta^{1/3}/\kappa)]. \tag{10}$$

If we allow k_1 to vary between 0 and ∞, then y varies from a lower value

$$y_0=R^2\beta^{-2/3};\quad R^2=(2\mu/\hbar^2)(E_g-\hbar\omega) \tag{11}$$

to ∞. Note that R^2 may be either positive or negative. We then find that the absorption coefficient is

$$\alpha=\frac{4\pi^2K\mu\beta^{2/3}}{\omega\kappa}\sum_{j=j_0}^{\infty}\mathrm{Ai}^2\left(\frac{2\pi j\beta^{1/3}}{\kappa}-\chi\right), \tag{12}$$

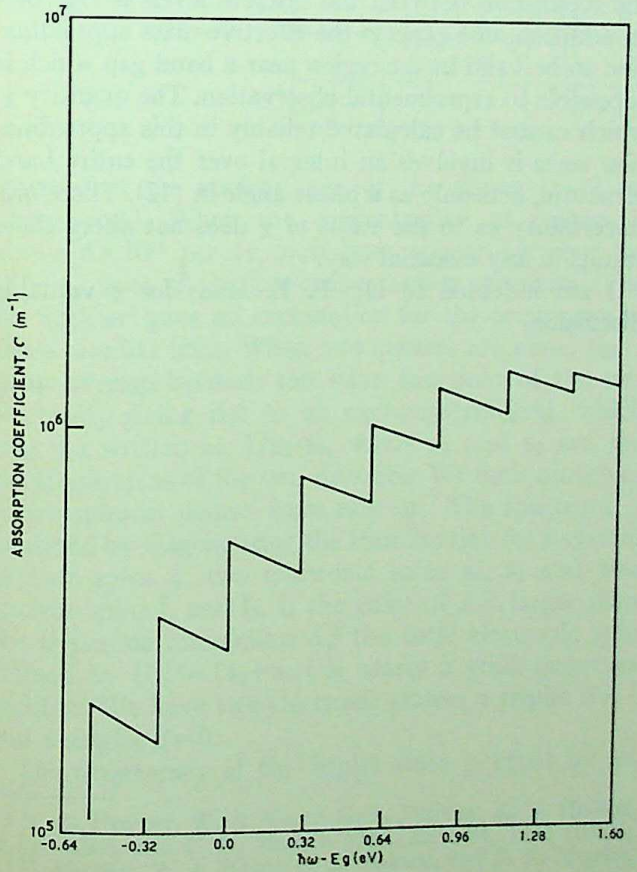


FIG. 1. The optical absorption coefficient α is shown as a function of $\hbar\omega-E_g$ as calculated from Eq. (12) using numbers appropriate to the transition between the heavy-hole band and the conduction band in gallium arsenide. For the purpose of illustration, a very large field, 10^9 eV/m has been chosen.

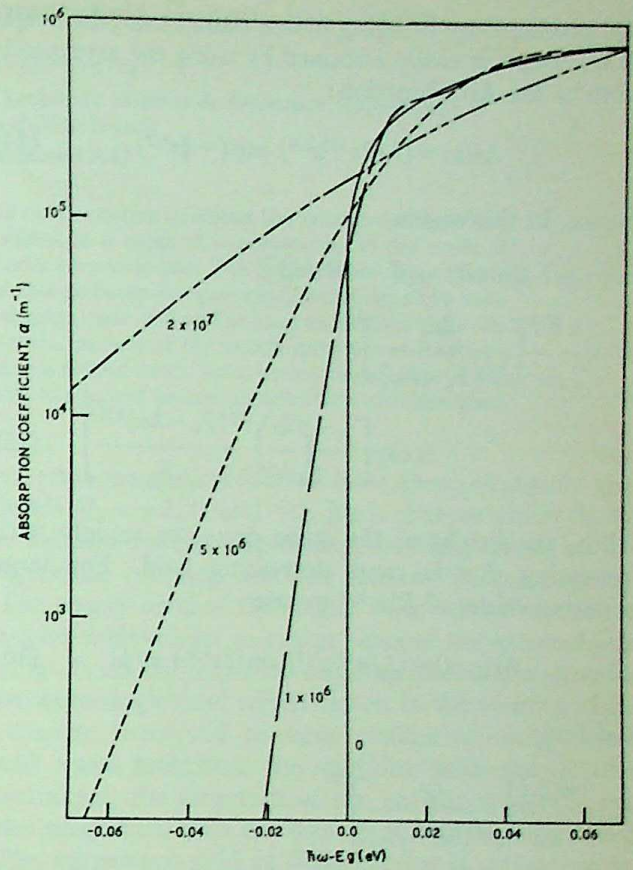


FIG. 2. The optical absorption coefficient is shown as a function of energy in the weak-field limit, Eq. (18), neglecting the discrete levels. Numbers appropriate to gallium arsenide, including both light and heavy holes, have been considered. Each curve is marked by the value of the field F in electron volts per meter which has been used in the calculation. The field is oriented along the [100] crystal axis.

in which the lower limit of the summation j_0 is determined by the condition that j_0 be the next integer larger than q_0, where

$$(2\pi q_0\beta^{1/3}/\kappa)-\chi=R^2/\beta^{2/3}. \tag{13}$$

Let us interpret this result. Observe that the argument of the Airy functions in (12) is independent of energy. The energy dependence of the absorption coefficient is contained in the external factor ω^{-1} and implicitly in the lower limit of the summation. As the photon energy $\hbar\omega$ increases, R^2 decreases, and consequently y_0 and j_0 decrease. For each decrease in j_0, a new term appears in the sum in Eq. (12). Consequently, a graph of the absorption coefficient α as a function of energy resembles a staircase with slanted steps. The width of each step is the energy required to produce a unit change in j_0. This is easily seen to be given by Eq. (5). Hence, the width of a step is just the separation of the discrete levels. The height of a step is proportional to $\mathrm{Ai}^2(R^2\beta^{-2/3})$, evaluated at the energy at which the jump occurs. The downward slant of each step is determined by the factor ω^{-1}.

For large positive values of $R^2\beta^{-2/3}$, the change in α,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$\Delta\alpha$, produced by bringing in one more term (the height of the step), is easily obtained by using the asymptotic form of the Airy function:

$$\mathrm{Ai}(z)=(1/2\pi^{1/2}z^{1/4})\exp(-\tfrac{2}{3}z^{3/2}). \tag{14}$$

Hence, in this region

$$\Delta\alpha\approx(\pi K\beta/\omega\kappa R)\exp[-4R^3/3\beta]$$

$$=\frac{\pi KF}{\kappa\omega}\left[\frac{2\mu}{\hbar^2(E_0-\hbar\omega)}\right]^{1/2}$$

$$\times\exp\left[-\frac{4}{3}\left(\frac{2\mu}{\hbar^2}\right)^{1/2}\frac{(E_0-\hbar\omega)^{3/2}}{F}\right]. \tag{15}$$

Thus, the height of the steps decreases rapidly with increasing $E_0-\hbar\omega$ and decreasing field. For large negative values of $R^2\beta^{-2/3}$ we use

$$\mathrm{Ai}(-z)=(1/\pi^{1/2}z^{1/4})\sin(\tfrac{2}{3}z^{3/2}+\pi/4) \tag{16}$$

and find

$$\Delta\alpha\approx\frac{4\pi KF}{\kappa\omega}\left[\frac{2\mu}{\hbar^2(\hbar\omega-E_0)}\right]^{1/2}$$

$$\times\sin^2\left[\frac{2}{3}\left(\frac{2\mu}{\hbar^2}\right)^{1/2}\frac{(\hbar\omega-E_0)^{3/2}}{F}+\frac{\pi}{4}\right]. \tag{17}$$

These results are illustrated in Fig. 1.

The mathematical argument may be concluded by observing that the weak-field limit of the absorption is obtained by replacing the sum in (12) by an integral. Put

$$z=2\pi j\beta^{1/3}/\kappa-\chi,$$

and let z be continuous. Then we obtain in place of (12)

$$\alpha=2\pi\omega^{-1}K\mu\beta^{1/3}\int_{y_0}^{\infty}\mathrm{Ai}^2(z)dz$$

$$=2\pi\omega^{-1}K\mu\beta^{1/3}\left[\left(\frac{d\,\mathrm{Ai}(y)}{dy}\right)_{y_0}^2-y_0\,\mathrm{Ai}^2(y_0)\right], \tag{18}$$

where y_0 is given in Eq. (11). This is equivalent to an expression given by Tharmalingam,[1] and may also be obtained by discarding all but the constant term in Eq. (1). Expressions valid in the limit of large positive or negative values of y_0 may be obtained by inserting the asymptotic expansions of the Airy functions Eqs. (14) and (16) into (18). The results of this can be obtained in an obvious way from Eqs. (23), (29), and (31) of I, and also from Tharmalingam's paper, so they will not be repeated here. However, since the weak-field absorption coefficient has not been exhibited in the region near the band gap in which asymptotic forms cannot be used, we present in Fig. 2 the absorption coefficient as calculated in this approximation using numbers appropriate to gallium arsenide.

Finally, we should note that although the detailed results of this calculation depend on the use of the effective-mass approximation, Eq. (4), the existence of structure in the absorption constant with the period of the separation between the discrete levels is rigorous. In addition, one expects the effective-mass approximation to be valid in the region near a band gap which is accessible to experimental observation. The quantity χ, which cannot be calculated reliably in this approximation since it involves an integral over the entire band structure, acts only as a phase angle in (12). Therefore, uncertainty as to the value of χ does not affect these results in any essential way.

I am indebted to Dr. P. E. Kaus for a valuable discussion.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Electron Spin Resonance on Interacting Donors in Silicon

D. JÉROME AND J. M. WINTER
Centre d'Etudes Nucléaires de Saclay, Service de Physique du Solide et de Résonance Magnétique, Gif-sur-Yvette, Seine-et-Oise, France

(Received 14 November 1963)

The value of the exchange integral J between pairs of donor centers in silicon has been investigated by looking at the ENDOR spectrum of phosphorus-doped silicon in a range of concentration of the order of 8×10^{16} per cc. The microwave field is saturating the cluster electronic line. The line shape of these ENDOR lines gives an unambiguous determination of the sign of this exchange integral which turns out to be antiferromagnetic. It is also possible to determine an order of magnitude of J and to have an idea about its distribution for two samples. We are able to understand all the features of the ENDOR spectrum at least in a qualitative way by calculating all the energy levels using a second-order perturbation calculation. A line coming from an ionized phosphorus center, weakly coupled to a pair of neutral phosphorus is also identified.

I. INTRODUCTION

THE substitutional donors like phosphorus or arsenic at low concentration in silicon give rise to a paramagnetic resonance spectrum[1,2] of $2I+1$ lines,

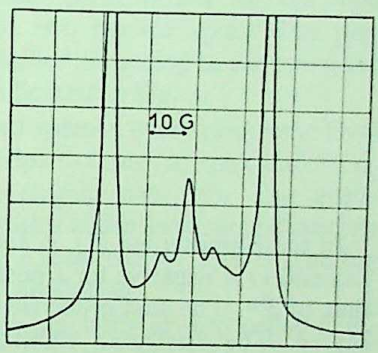


FIG. 1. Electron-resonance spectrum of phosphorus-doped silicon; 8×10^{16} P/cm³ at 1.3°K.

where I is the nuclear spin of the donor ($I=\frac{1}{2}$ for phosphorus). When the concentration of centers is above 6×10^{16} per cc, new lines appear between the $(2I+1)$ lines of the low-concentration spectrum (Fig. 1). Slichter[3] gave an explanation for the occurrence of these satellite lines. When two centers are close, there is an overlap between the wave functions of the two electrons, giving rise to an exchange integral, which may be written as $hJ\mathbf{s}_1\cdot\mathbf{s}_2$, where $\mathbf{s}_1$ and $\mathbf{s}_2$ are the electronic spins of the two electrons. We limit ourselves to phosphorus donors from now on. The spectrum is analyzed by diagonalizing the Hamiltonian for a system of four spins $\frac{1}{2}$, two electronic spins $\mathbf{s}_1$, $\mathbf{s}_2$ and two nuclear spins $\mathbf{I}_1$ and $\mathbf{I}_2$. If the value of J is larger than the hyperfine interaction A,[4] the total electronic spin defined by $|\mathbf{S}|^2=|\mathbf{s}_1+\mathbf{s}_2|^2$ is nearly a good quantum number. We have two electronic states; a triplet $S=1$ and a singlet $S=0$.

The degeneracy of the triplet state is lifted by an external magnetic field; we have three electronic sublevels $M_s=+1$, 0, and -1. Each of these sublevels has a fourfold nuclear degeneracy, partially lifted by the hyperfine coupling and the nuclear Zeeman energy. The energy level of the singlet state crosses one of the triplet sublevels for a certain value of the external field (Fig. 2). As the hyperfine coupling has matrix elements between singlet and triplet states, in the vicinity of the crossing, there will be large change of energy levels and wave functions. By studying with the ENDOR technique, the structure of the nuclear sublevels, and the magnitude and the sign of J will be determined. The microwave field at 9206 Mc/sec is saturating the cluster electronic line, therefore we select nuclear spins of pairs of exchange-coupled phosphorus nuclei (Fig. 1). In a given sample, there is a distribution of values of J due to the random distribution of phosphorus centers. We have been able to determine the value of J which has the largest probability for two samples with different phosphorus concentrations.

II. THEORY

1. First-Order Perturbation Theory of Energy Levels

We assume that the overlap of the wave functions belonging to two neighboring phosphorus atoms, leads

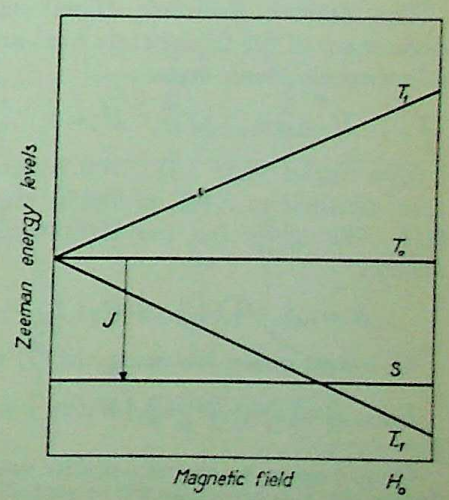


FIG. 2. Energy levels of the system of two electronic spins in the static magnetic field coupled by the exchange interaction.

[1] R. C. Fletcher, W. A. Yager, G. L. Pearson, A. N. Holden, W. T. Read, and F. R. Merritt, Phys. Rev. **94**, 1392 (1954); R. C. Fletcher, W. A. Yager, G. L. Pearson, and F. R. Merritt, Phys. Rev. **95**, 844 (1954).
[2] G. Feher, R. C. Fletcher, and E. A. Gere, Phys. Rev. **100**, 1784 (1955).
[3] C. P. Slichter, Phys. Rev. **99**, 479 (1955).
[4] A. Abragam (unpublished lectures).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

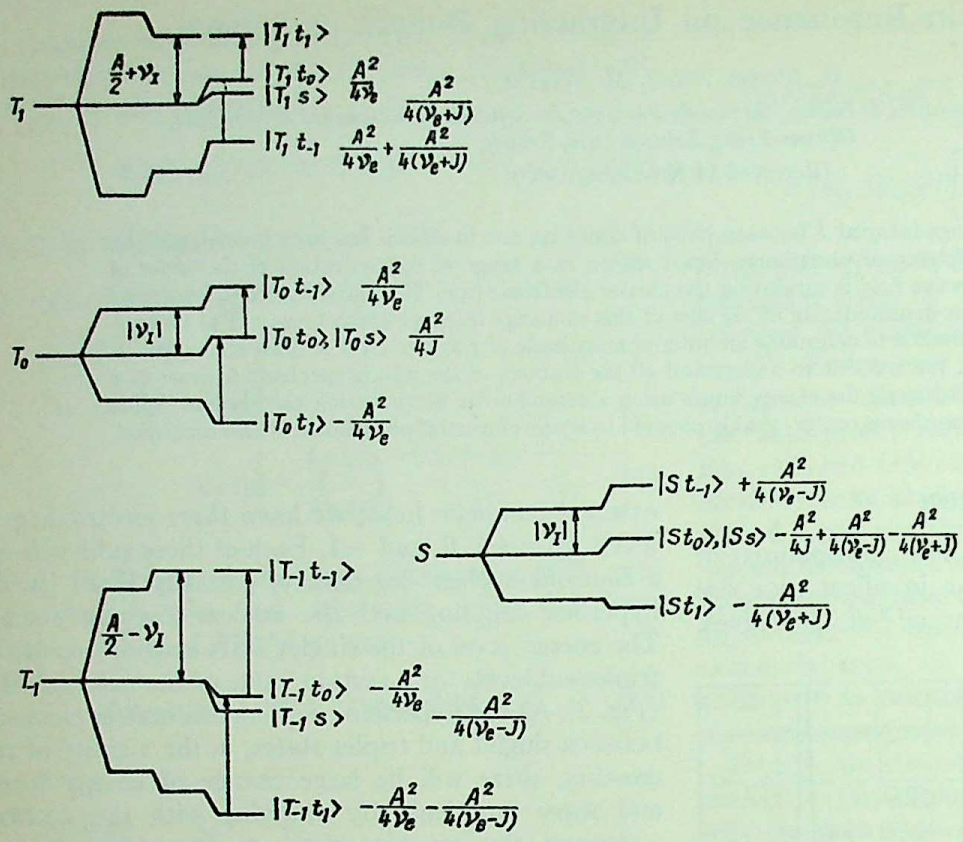


FIG. 3. Second-order shift of energy nuclear sublevels in the different triplet and singlet electronic states (not to scale). The good wave functions in zero-order approximation are noted beside each level. We assume an antiferromagnetic coupling lower than ν_e.

to a coupling between their spins $hJ\mathbf{s}_1\cdot\mathbf{s}_2$. The total Hamiltonian will be written in frequency units.[4]

$$\mathcal{H}=\nu_e(s_{1z}+s_{2z})+J\mathbf{s}_1\cdot\mathbf{s}_2+A(\mathbf{s}_1\cdot\mathbf{I}_1+\mathbf{s}_2\cdot\mathbf{I}_2)+\nu_I(I_{1z}+I_{2z}), \quad (1)$$

where ν_e and ν_I are the electronic and nuclear Larmor frequencies, respectively. J is the exchange coupling value which has been measured in cycles per second.[5]

To a first approximation, taking into consideration only the first two terms of the Hamiltonian (1), the total electronic spin is a good quantum number:

$$|\mathbf{S}|^2=|\mathbf{s}_1+\mathbf{s}_2|^2 \quad \text{and} \quad S_z=s_{1z}+s_{2z}$$

define good quantum numbers.

The Zeeman electronic Hamiltonian removes the degeneracy of the triplet state $S=1$ and gives the three following electronic states:

$$|T_1\rangle, \quad M_s=+1; \quad |T_0\rangle, \quad M_s=0; \quad |T_{-1}\rangle, \quad M_s=-1.$$

The singlet state $|S\rangle$, $S=0$ is diamagnetic. In the spin multiplicity $S=1$, $\mathbf{s}_1$ and $\mathbf{s}_2$ may be replaced by $\mathbf{S}/2$. This gives for the Hamiltonian of the whole system:

$$\mathcal{H}=\nu_e S_z+\tfrac{1}{4}J+\tfrac{1}{2}A\mathbf{S}\cdot(\mathbf{I}_1+\mathbf{I}_2)+\nu_I(I_{1z}+I_{2z}). \quad (2)$$

In lowest order, the energy of (2) will be given by

$$E_{M_s m_1 m_2}=\tfrac{1}{4}J+\nu_e M_s+\tfrac{1}{2}AM_s(m_1+m_2)+\nu_I(m_1+m_2).$$

[5] From now on, we assume a positive value for J. We shall see further that this hypothesis will be verified.

In the Hamiltonian (1), ν_e and A are positive quantities and ν_I is negative for a positive nuclear moment as that of P^{31}. The first-order energy levels are shown in Fig. 3. The electronic transition frequencies $\Delta S=0$, $\Delta M_s=\pm1$, $\Delta m_1=\Delta m_2=0$ are the following:

$$\nu=\nu_e+\tfrac{1}{2}A(m_1+m_2).$$

The electronic resonance spectra exhibit two lines at frequencies $\nu_e\pm\frac{1}{2}A$ which are not observable, due to the lines of the isolated donors and a line at the frequency ν_e twice as intense corresponding to $m_1+m_2=0$.[2,4]

We use the same considerations for clusters of three phosphorus atoms in the spin multiplicity $S=\frac{3}{2}$ which is completely symmetric with respect to the three spins $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$. We obtain for this case $\mathbf{s}_1=\mathbf{s}_2=\mathbf{s}_3=\frac{1}{3}\mathbf{S}$. The only electronic frequencies are $\nu_e\pm\frac{1}{2}A$ and $\nu_e\pm\frac{1}{6}A$.

The two eigenstates of spin $S=\frac{1}{2}$ give electronic resonance frequencies which vary from one cluster to another giving rise to a continuous background as has been shown by Abragam.[4]

The only discrete lines for four-atom clusters are ν_e, $\nu_e\pm\frac{1}{2}A$, $\nu_e\pm\frac{1}{4}A$. A line at frequency ν_e broadened by hyperfine interaction with Si^{29} spins appears only for clusters of an even number of phosphorus atoms.

At a temperature of 1.3°K the electronic resonance spectrum shows lines at frequencies $\nu_e\pm\frac{1}{4}A$ which merge almost completely in the background line. We conclude that the number of four-atom clusters is negligible in the samples (Fig. 1).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Allowed nuclear transitions and ENDOR frequencies.

$\lvert T_{-1},t_0\rangle \rightarrow \lvert T_{-1},t_{-1}\rangle$	$\frac{A}{2}-\nu_I+\frac{A^2}{4\nu_e}$	$\lvert T_1,t_0\rangle \rightarrow \lvert T_1,t_1\rangle$	$\frac{A}{2}+\nu_I-\frac{A^2}{4\nu_e}$	$\lvert T_0,t_1\rangle \rightarrow \lvert T_0,t_0\rangle$	$\lvert\nu_I\rvert+\frac{A^2}{4\nu_e}+\frac{A^2}{4J}$
$\lvert T_{-1},t_1\rangle \rightarrow \lvert T_{-1},t_0\rangle$	$\frac{A}{2}-\nu_I+\frac{A^2}{4(\nu_e-J)}$	$\lvert T_1,t_{-1}\rangle \rightarrow \lvert T_1,t_0\rangle$	$\frac{A}{2}+\nu_I-\frac{A^2}{4(\nu_e+J)}$	$\lvert T_0,t_0\rangle \rightarrow \lvert T_0,t_{-1}\rangle$	$\lvert\nu_I\rvert+\frac{A^2}{4\nu_e}-\frac{A^2}{4J}$

2. Second-Order Perturbation Theory of Energy Levels

There are nondiagonal matrix elements of the hyperfine structure Hamiltonian $\mathcal{H}_P=A(\mathbf{I}_1\cdot\mathbf{s}_1+\mathbf{I}_2\cdot\mathbf{s}_2)$ between different triplet states and between singlet and triplet states. The Oz component of the total spin $F_z=S_z+I_{1z}+I_{2z}$ and the total symmetry of wave functions with respect to interchange of two atoms are good quantum numbers. We defined a total parity number P as negative when the sign of the wave function changes by simultaneous permutation of the two electronic spins and the two nuclear spins. The perturbed states will be classified according to the two good quantum numbers as indicated in Fig. 3.

The coupling of nuclear spins in every electronic level gives three symmetric triplet states $\lvert t_1\rangle$, $\lvert t_0\rangle$, and $\lvert t_{-1}\rangle$, and an antisymmetric singlet state $\lvert s\rangle$. The even-interaction $\mathcal{H}_P$ couples states which have same quantum numbers F_z and the same total parity. This is true for any order perturbation calculation.

Starting with the eigenfunctions of zero order (Fig. 3), we obtain immediately the nondiagonal matrix elements from which one can obtain the shift of energy levels.

We present the calculation of the shift for triplet states $\lvert T_1\rangle$ and $\lvert T_0\rangle$. We deduce shifts of $\lvert T_{-1}\rangle$ levels by exchanging ν_e in $-\nu_e$ in the formula.

(a) $\lvert T_1,t_1\rangle$. This state is not coupled with any other state triplet or singlet because $\lvert T_1,t_1\rangle$ is the only state with $F_z=2$ and there is no change in its energy beyond the first-order value.

(b) $\lvert T_1,t_0\rangle$ *and* $\lvert T_1,s\rangle$. $\lvert T_1,t_0\rangle$ is coupled to $\lvert T_0,t_0\rangle$ and $\lvert T_1,s\rangle$ to $\lvert S,t_1\rangle$ by the nondiagonal matrix elements. There are no diagonal matrix elements for any of these states,

$$\langle T_1,t_0\lvert\mathcal{H}_p\rvert T_0,t_1\rangle=A/2,$$
$$\langle T_1,s\lvert\mathcal{H}_p\rvert S,t_1\rangle=-A/2.$$

The second-order shift of the state $\lvert T_1,s\rangle$ is

$$E=A^2/4(\nu_e+J),$$

and that of state $\lvert T_1,t_0\rangle$ is

$$E=A^2/4\nu_e.$$

(c) $\lvert T_1,t_{-1}\rangle$. This state is coupled to $\lvert T_0,t_0\rangle$ and $\lvert S,s\rangle$. Its shift is

$$A^2/4\nu_e+A^2/4(\nu_e+J).$$

(d) $\lvert T_0,t_0\rangle$ *and* $\lvert T_0,s\rangle$. The same method as that of Sec. 2(b) does not remove the degeneracy between the two states in second order. To that order both states are shifted by the same amount,

$$E=A^2/4J.$$

(e) $\lvert T_0,t_1\rangle$ *and* $\lvert T_0,t_{-1}\rangle$. The shifts are respectively $A^2/4\nu_e$ and $-A^2/4\nu_e$. This computation is only valid under conditions: $A\ll\nu_e$, J, ν_e-J, ν_e+J. It is not valid at the crossing of energy levels. This restriction is important as we shall see further.

3. Line Shapes and Frequencies of Nuclear Transitions

(a) The coupling with the rf field inducing an electronic or nuclear transition is an operator proportional to

$$\alpha(s_{1z}+s_{2z})+\beta(I_{1z}+I_{2z}).$$

The selection rules for electronic or nuclear transitions are thus $\Delta F_z=\pm1$, $\Delta P=0$. With a perturbation like $s_{1z}+s_{2z}$ or $I_{1z}+I_{2z}$ the total parity is preserved. The allowed transitions are shown in Table I and Fig. 3.

(b) From Table I, it appears that nuclear lines are

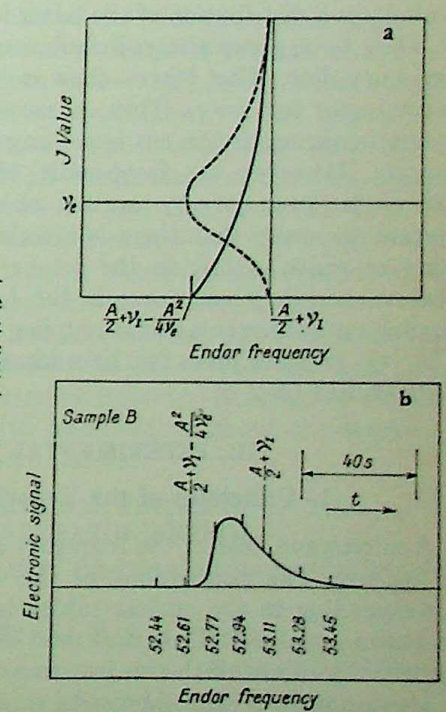


FIG. 4. (a)–Theoretical dependence on J of the ENDOR frequency in the group $\frac{1}{2}A+\nu_I$ and estimate of J distribution in dashed line. (b)–Experimental line shape for sample B.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

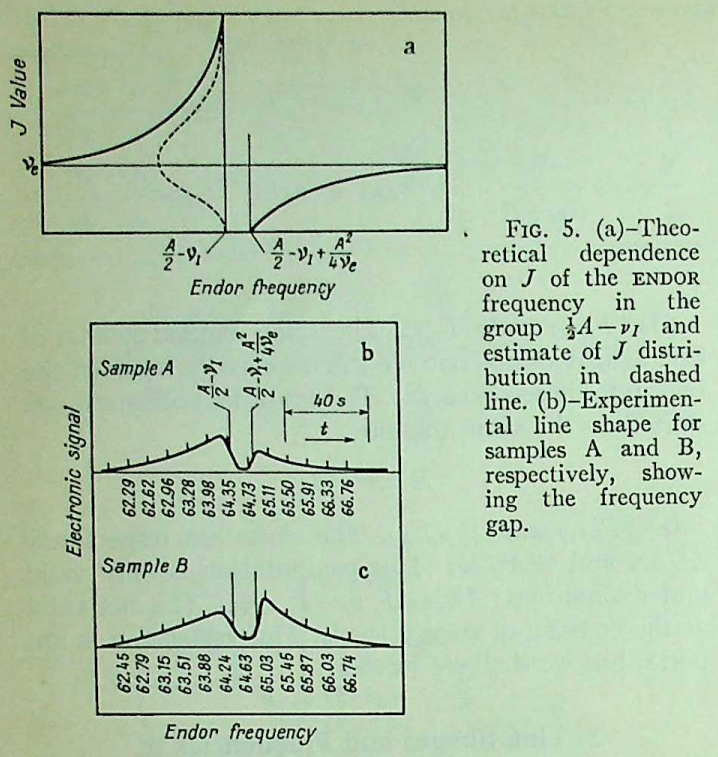


FIG. 5. (a)–Theoretical dependence on J of the ENDOR frequency in the group $\frac{1}{2}A-\nu_I$ and estimate of J distribution in dashed line. (b)–Experimental line shape for samples A and B, respectively, showing the frequency gap.

split into three groups, two whose frequencies in first order are $\frac{1}{2}A-\nu_I$ and $\frac{1}{2}A+\nu_I$, and one whose mean frequency is near the Larmor frequency of a bare phosphorus atom. We assume now $J>0$. The values of the coupling constant are distributed in the clusters with a distribution function $P(J)$ normalized to unity when J is varied from zero to infinity. It is assumed to have a maximum for a value $J=J_P$.

The $\frac{1}{2}A+\nu_I$ group includes two frequencies. One is J-independent $\frac{1}{2}A+\nu_I-(A^2/4\nu_e)$, the other is J-dependent $\frac{1}{2}A+\nu_I-[A^2/4(\nu_e+J)]$. The distribution of J produces a distribution of the latter frequencies (Fig. 4). The $\frac{1}{2}A-\nu_I$ group always includes one J-independent frequency line. The others show a singular energy denominator for $J=\nu_e$. Thus, a second-order perturbation-frequency treatment is no longer valid for such clusters. However the frequencies of these clusters ($J\simeq\nu_e$) far from $\frac{1}{2}A-\nu_I$, are not observed. It is important to notice that there is a forbidden frequency band of width $A^2/4\nu_e$ in the $\frac{1}{2}A-\nu_I$ group. Such a phenomenon does not occur in the $\frac{1}{2}A+\nu_I$ group assuming an antiferromagnetic coupling (Figs. 4 and 5). The $|\nu_I|$ group includes two lines whose frequencies are J-dependent (Fig. 6).

III. EXPERIMENTAL

1. Principles of the Experiment

A microwave field at the frequency $\nu_e=9206$ Mc/sec is applied. The populations of the electronic levels corresponding to the nuclear sublevels $m_1+m_2=0$ are therefore equalized. Then a rf field of frequency ν in the vicinity of one of the nuclear-transition frequencies is also applied. This field produces a change in the population of one of the states with $m_1+m_2=0$ and the electronic signal varies. By sweeping the radio-frequency ν we reach all the nuclear frequencies possible.[6] For a given ν the change of the electronic signal is very likely proportional to the number of clusters having the nuclear frequency transition ν. For instance, if we look at the ENDOR frequencies $\nu_I+(A^2/4\nu_e)\pm(A^2/4J)$, the spectrum reproduces the distribution of the values of $A^2/4J$ in our sample.

2. Experimental Technique

The frequency of the electronic spectrometer is stabilized by a Pound discriminator on an external cavity. We use a constant-level symmetric detection.[7] All these experiments have been done at 1.3°K and because of the length of the electronic relaxation time, the dispersion mode is observed.[6] The signal-to-noise ratio for the ENDOR signal on clusters varies between 3 and 5. We use two samples of phosphorus-doped silicon from Merck;

Sample A: $N=8\times10^{16}\,P/\text{cc}$ $\rho=0.11\Omega$-cm,

Sample B: $N=6\times10^{16}\,P/\text{cc}$ $\rho=0.13\Omega$-cm.

3. Experimental Results

Taking for the phosphorus nuclear Larmor frequency the value $|\nu_I|=5.67$ Mc/sec and for the hyperfine constant the value[6] $A=117.53$ Mc/sec, the nuclear-frequency transition calculated in a first-order approximation is given by $\frac{1}{2}A-\nu_I=64.11$ Mc/sec, $\frac{1}{2}A+\nu_I=53.09$ Mc/sec, and $|\nu_I|$. A typical second-order correction is given by $A^2/4\nu_e=0.375$ Mc/sec.

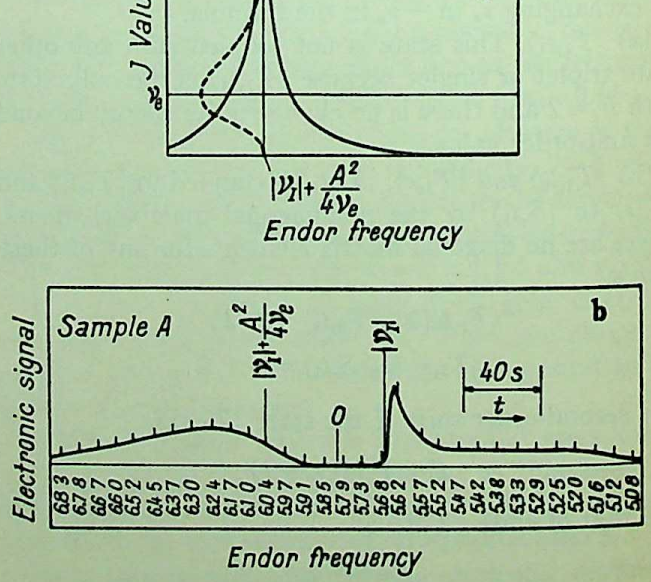


FIG. 6. (a)–Theoretical dependence on J of the ENDOR frequency in the group $|\nu_I|$. (b)–Experimental line shape for sample A.

[6] G. Feher, Phys. Rev. 114, 1219 (1959).
[7] G. Feher, Bell System Tech. J. 26, 449 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The ENDOR line in the vicinity of $\nu=\frac{1}{2}A-\nu_I$ shows two peaks separated by a frequency gap of the order of $A^2/4\nu_e$. This fact shows without any ambiguity that the singlet level crosses the group of $M_s=-1$ sublevels and the coupling is antiferromagnetic (Fig. 5).

The line for $M_s=1$ sublevels shows only a broadening (Fig. 4) coming from the different possible value of the second-order shift $A^2/4(\nu_e+J)$.

The ENDOR line for the $M_s=0$ group of levels (Fig. 6) occurs at the frequencies ν given by $\nu=|\nu_I|+(A^2/4\nu_e)\pm(A^2/4J)$.

As stated above, the shape of the line reproduces the distribution of $A^2/4J$. If we call J_F the value of J which gives the maximum probability for $A^2/4J$, we deduce $A^2/2J_F$ by measuring the distance between the maximum of the two peaks. For the sample A we find

$$A^2/2J_F=0.94 \text{ Mc/sec}.$$

Knowing the distribution for $A^2/4J$, it is possible to deduce the function $P(J)$, giving the probability of finding a given J and in particular, the value J_P giving the maximum value for $P(J)$. The results are

$$\text{Sample A:}\quad J_P=15\,000\pm3\,000 \text{ Mc/sec},$$

$$\text{Sample B:}\quad J_P=13\,500\pm3\,000 \text{ Mc/sec}.$$

The shape of the two other lines can be discussed in a qualitative way. Figure 5 shows the lines $\frac{1}{2}A-\nu_I$ for the two samples; for the sample A the low-frequency peak is more intense than the high-frequency peak. The result is opposite for sample B.

Since the low-frequency peak corresponds to values of J larger than ν_e and the other peak to values of J smaller than ν_e, this result is consistent with the fact that J_P is smaller in sample B than in sample A (Fig. 5).

The width of the experimental frequency gap is smaller than $A^2/4\nu_e$; the time constant due to electronic saturation is probably reducing the experimental gap.

Coming back to the line at ν_I we notice (Fig. 6) the two broad lines. The intensity of these lines is very small in sample B. The width of the line, being of the order of $A^2/4J_P$, is larger for sample B and is responsible for the small signal.

IV. INFLUENCE OF IONIZED CENTERS

A very narrow line (10 kc/sec) occurring exactly at the frequency $|\nu_I|$ is observed in the ENDOR spectrum of both samples (Fig. 6). Feher[6] already noticed the occurrence of this line in the ordinary ENDOR spectrum of a compensated sample and attributed it to an ionized phosphorus center.

Assuming that there is an ionized phosphorus in the vicinity of a cluster of two neutral phosphorus, there always will be a small scalar coupling between the two electronic spins and the ionized phosphorus nuclear spin $\mathbf{I}'$. We therefore add, to our Hamiltonian (1), the term

$$a_1\mathbf{I}'\cdot\mathbf{s}_1+a_2\mathbf{I}'\cdot\mathbf{s}_2+\nu_I I_z'.$$

This coupling comes because the wave function for the two electrons localized around phosphorus 1 and 2, has a small admixture of a wave function for an electron bound to the phosphorus I'. This is an extension to a system of three phosphorus and two electrons of the calculation performed by Miller and Abrahams[8] for explaining the conduction of compensated sample. Because the admixture is very small a_1 and a_2 are much smaller than A. This new term gives a contribution to the energy spectrum, which is for a state M_s (of the triplet):

$$\nu_I m'+\tfrac{1}{2}(a_1+a_2)m'M_s,$$

m' being the value of I_z'.

As a_1 and a_2 are small this term does not influence the electronic spectrum. But there are now ENDOR lines corresponding to the selection rule $\Delta m'=\pm1$, and occurring at frequencies given by

$$\nu=|\nu_I|+\tfrac{1}{2}(a_1+a_2)M_s$$

(we neglect second-order corrections). For $M_s=0$ a line appears exactly at $\nu=|\nu_I|$.

There are also ENDOR lines corresponding to value of $M_s=\pm1$, but they are not observed because there is a distribution of a_1 and a_2 values.

The important point is to notice that only lines corresponding to the $M_s=0$ level of clusters are observed. There is undoubtedly phosphorus ionized near isolated neutral phosphorus, but their ENDOR frequencies are given by $\nu=|\nu_I|+\frac{1}{2}aM_s$, $M_s=\pm\frac{1}{2}$ and are not observed. We observe only ionized phosphorus coupled to clusters of two neutral phosphorus, this point is also proved by the following experimental fact.

Now we saturate the electronic line $\nu=\nu_e\pm\frac{1}{2}A$ (Fig. 1) and look at the ENDOR spectrum. The narrow line at $|\nu_I|$ appears but its intensity is half the intensity of the $|\nu_I|$ line observed when we saturate the $\nu=\nu_e$ line. It is known that a cluster of two has an electron spectrum with lines at $\nu=\nu_e\pm\frac{1}{2}A$ which are not observed because they are masked by the large signal coming from isolated phosphorus, the intensity of these lines is one-half the intensity of the $\nu=\nu_e$ central line. So the narrow $\nu=|\nu_I|$ line observed when we saturate the electronic lines at $\nu=\nu_e\pm\frac{1}{2}A$ is not due to the ENDOR spectrum of isolated phosphorus but to the ENDOR spectrum of clusters of phosphorus.

By comparing the intensity of the narrow $\nu=|\nu_I|$ line and the intensity of the broad line near $|\nu_I|$ (coming from neutral phosphorus) it should be possible to get quantitative informations about the degree of compensation of the samples.

V. THEORETICAL ESTIMATE OF THE EXCHANGE ENERGY

The Hamiltonian of the two-electron system is given by

$$\mathcal{H}(1,2)=T(\mathbf{r}_1)+T(\mathbf{r}_2)+V_i(\mathbf{r}_1)+V_j(\mathbf{r}_1)+V_i(\mathbf{r}_2)+V_j(\mathbf{r}_2)+V(\mathbf{r}_1,\mathbf{r}_2),$$

[8] A. Miller and E. Abrahams, Phys. Rev. 120, 745 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

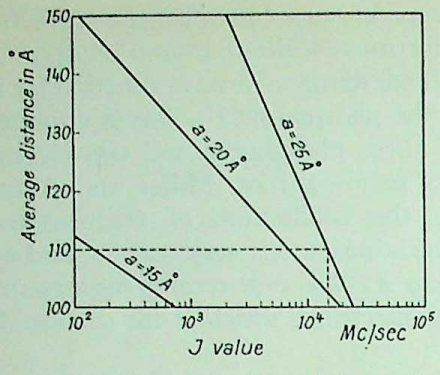


FIG. 7. Results of the computation of J as a function of the average distance between two neighboring donors. The experimental value of J for sample A corresponds to an average distance of 110 Å, assuming $a=25$ Å.

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the coordinates of the two electrons, $T(\mathbf{r})$ is the kinetic-energy operator, $V_i(\mathbf{r})$ and $V_j(\mathbf{r})$ are the attractive potentials due to phosphorus nuclei at sites i and j. In our spin Hamiltonian (1), J is defined as the energy difference between triplet and singlet spin states[9,10]

$$J=\langle T|\mathcal{H}(1,2)|T\rangle-\langle S|\mathcal{H}(1,2)|S\rangle,$$

where $|T\rangle$ is the total wave function for the system, antisymmetric with respect to the orbital variables, symmetric with respect to the spin variables in the triplet spin state. The opposite is true for $|S\rangle$.

The orbital part of the one-electron wave function, taken from the effective-mass theory[11] may be written at a site i:

$$\psi_i(\mathbf{r})=\sum_{p=1}^{6}(6)^{-1/2}F_p(\mathbf{r})\phi_p(\mathbf{r}),$$

where $\phi_p(\mathbf{r})$ is a Bloch function corresponding to the wave vector of the p valley, and $F_p(\mathbf{r})$ is an envelope function satisfying the wave equation in the effective-mass approximation:

$$F_p(\mathbf{r})=(\pi a^2 b)^{-1/2}\exp-\left[\frac{x^2+y^2}{a^2}+\frac{z^2}{b^2}\right]^{1/2},$$

the z axis being parallel to the wave vector corresponding to the value p. In silicon a and b are[12]

$$a=25\text{ Å},\quad b=14.2\text{ Å}.$$

According to our definition of the exchange energy, J takes the following form[9]:

$$J=-2\int\psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\int\psi_j^*(\mathbf{r}_2)V(\mathbf{r}_2)\psi_i(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$
$$-\int\psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\frac{e^2}{\kappa_0}|\mathbf{r}_1-\mathbf{r}_2|^{-1}\psi_j^*(\mathbf{r}_2)\psi_i(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2.$$

Since $V(\mathbf{r}_2)$ is an attractive potential the first term is a positive one, the second is negative. If one neglects the anisotropy of the envelope function within a given valley, the estimate of J is simple. But one should notice that each term of the latter expression is smaller than the one obtained by using a pure hydrogenic wave function because there is destructive interference between parts of the wave functions coming from different valleys.

Utilizing Eq. (II 17), Eq. (II 18), and (paragraph) IV of Ref. 8, we evaluated J (as function of the distance between phosphorus centers) assuming an isotropic envelope function and taking for the radius $a=25$, 20, or 15 Å (Fig. 7). We find that the positive theoretical value of J agrees with the experimental results. The average distance between centers is easily deduced from the concentration of centers, and we find that the experimental value of J is larger than the theoretical prediction, even if we use $a=25$ Å for the radius. If we believe in the impurity concentration measurements by conductivity at room temperature, this fact seems to indicate that the average distance between donors is smaller than the average distance deduced from the concentration, by about 20%. This may be due to a nonuniform concentration of phosphorus donors. It has been suggested that impurities tend to concentrate along dislocation lines.[13]

VI. CONCLUSION

The method we described can only give information if the exchange integral is larger than the hyperfine coupling A, otherwise no electronic line at the frequency ν_e appears in the spectrum. There is also an upper limit because when J becomes larger than kT and $h\nu_e$; there are very few electrons in the triplet states. If we remark that J varies very rapidly with the distance, our method is only valid in a narrow range of concentration. At high concentration, another limit appears. The electronic relaxation time becomes shorter, and the line difficult to saturate.

The possibility of measuring J is directly related to the presence of the hyperfine coupling term; without it, the exchange term and the Zeeman term commute and the frequencies are not affected by the exchange term. It is the small coupling by the hyperfine energy of the triplet and singlet levels which allows measurement of J. The effect of exchange on the energy spectrum is very small except when the crossing occurs. By a second-order perturbation calculation we are able to explain the shape of the experimental ENDOR spectrum and to determine the sign and the magnitude of J.

We have explained the occurrence of an ENDOR line

[9] P. W. Anderson, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Vol. 14.
[10] W. Heitler and F. London, Z. Physik 44, 455 (1927).
[11] J. M. Luttinger and W. Kohn, Phys. Rev. 97, 869 (1955).
[12] W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5.
[13] D. K. Wilson, Ph.D. dissertation, Rutgers University, New Brunswick, New Jersey, May 1963 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

coming from an ionized phosphorus center weakly coupled to a cluster of two neutral phosphorus atoms.

Note added in proof. Dr. J. J. Pearson made a calculation of J without neglecting the anisotropy of the envelope function. The results will be published in a separate article.

ACKNOWLEDGMENTS

The authors are especially indebted to Professor A. Abragam for very helpful suggestions, and to Dr. Ch. Ryter and Dr. C. Robert for their help on the experimental part of this work.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Peierls Stress and Creep of a Linear Chain*

J. H. WEINER AND W. T. SANDERS
Department of Mechanical Engineering, Columbia University, New York, New York
(Received 19 December 1963; revised manuscript received 23 January 1964)

The Frenkel-Kontorova dislocation model is modified by replacing the sinusoidal substrate force by one which is piecewise linear. Exact solutions are found for the static configuration and for the Peierls stress, σ_P. Good agreement is found between these values of σ_P and those obtained previously for a two-dimensional Rosenstock–Newell model. The atoms of the linear chain are then considered in random motion corresponding to thermal equilibrium and under an applied stress $\sigma<\sigma_P$. The time required for motion of the dislocation from one position of stable equilibrium to an adjacent one is computed by means of a rate-theory formulation adapted to the present type of problem in which the positions of all the atoms in the chain are required to vary in passing over the potential barrier. The theoretical transition times for an infinite chain are compared with analog computer results for a six-atom chain and reasonably good agreement is found.

1. INTRODUCTION

THE rate of dislocation creep over the Peierls barrier has been studied in connection both with low-temperature creep[1] and the Bordoni peak.[2] Because of the complexity of the phenomenon, it is necessary to make various assumptions of a mathematical and physical nature when treating a realistic model of the process.

The purpose of the present work has been to construct a simple dislocation model which is amenable to analysis with few additional assumptions. It is hoped that the results of this idealized analysis may provide insight into the nature of the real process, and that the mathematical techniques employed may be applied to more realistic models.

The dislocation model considered here is a modification of the Frenkel-Kontorova one-dimensional model[3] with the sinusoidal substrate force replaced by one which is piecewise linear as in the two-dimensional treatment of Sanders.[4] For this model it is possible to obtain exact solutions to the difference equation describing the static configuration under applied stress and to derive an exact expression for the Peierls stress, σ_P (Sec. 2). The values of σ_P computed here agree surprisingly well with the results of the two-dimensional calculations.[4]

In order to study the dislocation creep rate, the atoms of the model are then considered in random motion corresponding to thermal equilibrium and under an applied stress $\sigma<\sigma_P$. The problem of determining the rate at which the dislocation passes from one stable equilibrium position to an adjacent one is in the general class of problems considered in rate theory.[5] We present here (Sec. 3) a derivation of the pertinent rate formula *ab initio* which is somewhat different from the usual one and is particularly adapted to the present type of problem in which the positions of all the atoms in the chain are required to vary in passing over the potential barrier.

For the infinite chain dislocation model it is possible (Sec. 4) to calculate explicitly all the quantities entering into this rate formula. The theoretically predicted values of the frequency of transition of the dislocation from one equilibrium position to an adjacent one are found to be in order-of-magnitude agreement with analog

* This research was supported jointly by the U. S. Air Force Office of Scientific Research under Grant No. AF-AFOSR-228-63 and by the National Science Foundation under Grant No. NSF-G19010.

[1] J. Lothe and J. P. Hirth, Phys. Rev. 115, 543 (1959).

[2] A. Seeger, H. Donth, and F. Pfaff, Discussions Faraday Soc. 23, 19 (1957).

[3] J. Frenkel and T. Kontorova, Phys. Z. Sowjetunion 13, 1 (1938); J. Phys. (U.S.S.R.) 1, 137 (1939). A paper which appeared after this manuscript was submitted [J. Kratochvil and V. L. Indenbom, Czech. J. Phys. 13, 814 (1963)] treats the same modification of the Frenkel-Kontorova model from the static viewpoint only.

[4] W. T. Sanders, Phys. Rev. 128, 1540 (1962).

[5] H. Eyring, J. Chem. Phys. 3, 107 (1935); C. Wert and C. Zener, Phys. Rev. 76, 1169 (1949); C. Wert, *ibid.* 79, 601 (1950). A paper which came to our attention after this manuscript was submitted is G. H. Vineyard, Phys. Chem. Solids 3, 121 (1957), which also develops the pertinent rate formula from the viewpoint of Sec. 3. We have retained the present discussion, which differs somewhat in emphasis, for the sake of completeness.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

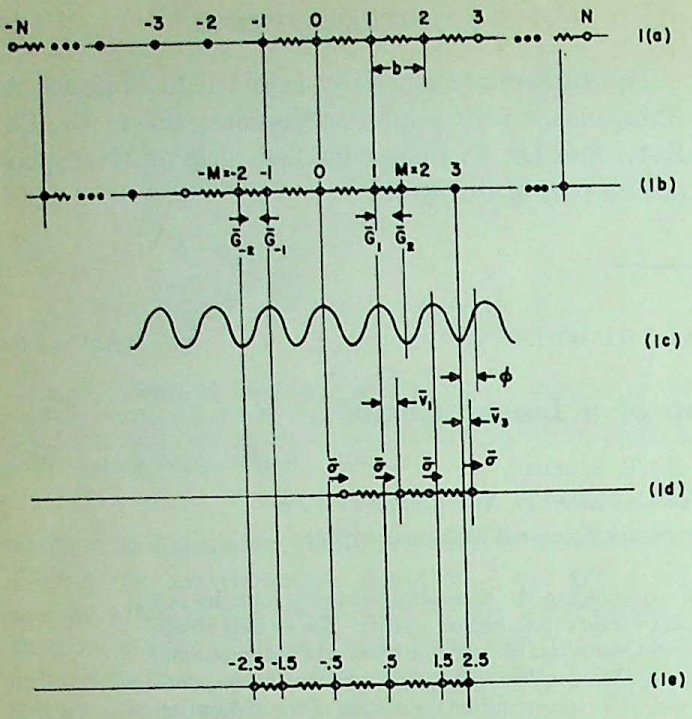


FIG. 1. Linear chain dislocation model.

computer results for a 6-atom chain. Finally, a summary and conclusions are presented in Sec. 5.

2. STATIC SOLUTION

Model Description

As a simplified crystal model, we consider a linear chain of mass points (hereafter referred to as atoms) interconnected by linear springs with spring constant k_1 and equilibrium spacing b (corresponding to the lattice parameter) and subjected also to a periodic (with period b) substrate potential. This substrate potential, as in the original Frenkel-Kontorova model, represents the net effect of the remaining atoms in the (3-dimensional) crystal. We assume a piecewise quadratic, continuous potential $U(x)$, namely,

$$\begin{aligned} U(x)&=\tfrac{1}{2}k_2x^2, && |x|\leq\phi,\\ &=\tfrac{1}{4}k_2\phi b-\frac{k_2\phi}{b-2\phi}\left(\frac{b}{2}-x\right)^2, && \phi\leq|x|\leq\frac{b}{2}, \end{aligned} \tag{2.1}$$

where x is the distance from a potential minimum. This replacement of the trigonometric potential of Frenkel and Kontorova by the piecewise quadratic potential of Eq. (2.2a) simplifies the analysis and introduces the additional parameter ϕ, the distance from a potential minimum to the nearest point at which the potential curvature is discontinuous and becomes negative; ϕ is thus a measure of the hardness of the crystal. The parameter k_2 in Eq. (2.1) is a measure of the shear modulus of the crystal while the parameter k_1 is a measure of the Young's modulus of the crystal.

In order to introduce a dislocation into this model, we start with the linear chain as in Fig. 1(a) and subject the jth atom to a longitudinal force $\bar{G}_j$, where

$$\bar{G}_M=\bar{G}_{-M+1}=-\tfrac{1}{2}k_1b;\quad \bar{G}_{M-1}=\bar{G}_{-M}=\tfrac{1}{2}k_1b;$$
$$\bar{G}_j=0 \quad \text{all other } j. \tag{2.2a}$$

From this, the following equilibrium configuration will result: All atoms $j\geq M$ are displaced $b/2$ to the left, all atoms $j\leq -M$ are displaced $b/2$ to the right, and those for which $-M<j<M$ are undisturbed.[6] To verify that this is indeed an equilibrium configuration, observe that the sum of the forces acting on each atom vanishes. For the special case $M=2$, this new equilibrium configuration is shown in Fig. 1(b). This configuration provides a model for a dislocation because five atoms, viz., -2, -1, 0, 1, and 2, now occupy a region ordinarily occupied by only four atoms.

Now with these forces still applied, the atoms are subjected to the periodic substrate potential defined in Eq. (2.1) and sketched in Fig. 1(c), and which is positioned relative to the configuration of Fig. 1(b) with either a maximum or minimum at the position of all atoms so that it exerts no forces on them in this configuration.

We next apply forces $-\bar{G}_j$ to the configuration of 1(b) and (c), annulling those originally applied, and also apply a constant force $\bar{\sigma}$ (which represents externally applied shear stress) to each atom. Let the resulting displacement of the jth atom from the configuration of 1(b) be $\bar{v}_j$. With the substrate potential of Eq. (2.1), the substrate force on the jth atom is

$$\begin{aligned} &F_j=-k_2\bar{v}_j \quad \text{for} \quad |j|\geq M, \quad \text{providing} \quad |\bar{v}_j|\leq\phi,\\ &F_j=\frac{2k_2\phi}{b-2\phi}\bar{v}_j \quad \text{for} \quad |j|<M,\\ &\qquad\qquad \text{providing} \quad |\bar{v}_j|\leq\left(\frac{b}{2}-\phi\right). \end{aligned} \tag{2.2b}$$

Note that in the first of Eqs. (2.2b) $\bar{v}_j$ is measured from a potential minimum, while in the second of these equations $\bar{v}_j$ is measured from a potential peak. As will be seen subsequently, the value of M is determined for a given set of parameters so that the inequalities on the atom displacements associated with Eq. (2.2b) are satisfied. The final configuration, after forces $-\bar{G}_j$ and $\bar{\sigma}$ are applied is then as shown in Fig. 1(d). The atoms in the range $|j|<M$ will be termed subject to weak bonds to the substrate potential while the others will be referred to as strong bonds. The number of weak bonds, $n_w=2M-1$, may be odd or even. In the latter case, the atoms will be indexed as shown in Fig. 1(e); that is, j is understood to take on the values $\cdots, -1\tfrac{1}{2}, -\tfrac{1}{2}, +\tfrac{1}{2},$

[6] Only the infinite chain will be considered in detail and therefore the location of the origin may always be taken as the midpoint of the dislocation. The indices j and M in Eq. (2.2a) may be either integers or half-odd integers; see discussion following Eq. (2.2b).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$1\frac{1}{2}, \cdots$ when $2M$ is odd. Therefore in both cases the weak-bond atoms may be taken to correspond to $-M<j<M$ and both cases may be treated together. The equations of equilibrium for the configuration of 1(d) are

$$k_1(\bar{v}_{j+1}+\bar{v}_{j-1}-2\bar{v}_j)-k_2\bar{v}_j-\bar{G}_j+\bar{\sigma}=0\,,\quad |j|\geq M\,,$$

$$k_1(\bar{v}_{j+1}+\bar{v}_{j-1}-2\bar{v}_j)+\frac{2k_2\phi}{b-2\phi}\bar{v}_j-\bar{G}_j+\bar{\sigma}=0\,, \tag{2.3}$$

$$-M<j<M\,.$$

With the introduction of the dimensionless variables

$$\begin{aligned} &v_j=\bar{v}_j/b\,, && G_j=\bar{G}_j/(k_1b);\\ &\sigma=\bar{\sigma}/(k_2b)\,, && P=k_2/k_1;\\ &\gamma=\phi/b\,, && Q=\frac{2\gamma P}{1-2\gamma}; \end{aligned} \tag{2.4a}$$

and the reduced displacements

$$u_j=v_j-\sigma\,, \tag{2.4b}$$

these equations may be written

$$u_{j+1}-(2+P)u_j+u_{j-1}=G_j\,,\qquad |j|\geq M; \tag{2.5a}$$

$$u_{j+1}-(2-Q)u_j+u_{j-1}=G_j-\frac{\sigma P}{(1-2\gamma)},$$

$$-M<j<M\,. \tag{2.5b}$$

Displacement Solution

We treat here only the infinite chain, $N\to\infty$; the solutions to Eq. (2.5) which decay at infinity are

$$\begin{aligned} &u_j=A\beta^{-j}, && j\leq -M;\\ &u_j=B\beta^{j}, && j\geq +M;\\ &u_{M-1}=B\beta^{M-1}-\tfrac{1}{2}, && u_{-M+1}=A\beta^{M-1}+\tfrac{1}{2};\\ &u_j=C\cos j\theta+D\sin j\theta-\frac{\sigma}{2\gamma}, && -M<j<M;\\ &u_M=C\cos M\theta+D\sin M\theta-\frac{\sigma}{2\gamma}+\tfrac{1}{2};\\ &u_{-M}=C\cos M\theta-D\sin M\theta-\frac{\sigma}{2\gamma}-\tfrac{1}{2}; \end{aligned} \tag{2.6a}$$

where

$$\beta=\tfrac{1}{2}[P+2-(P^2+4P)^{1/2}], \tag{2.7}$$

$$\cos\theta=1-Q/2\,.$$

Equations (2.6a) and (2.6b) satisfy (2.5a) and (2.5b), respectively. The four constants A, B, C, and D may then be determined by the requirement that Eqs. (2.6a) and (2.6b) agree for the four atoms, $-M$, $-M+1$, $M-1$, M for which the expressions overlap. The resulting expressions for the displacements are

$$\begin{aligned} u_j(\sigma;M)&=(\sigma F_M+G_M)\beta^{-j}; && j\leq -M;\\ &=(\sigma F_M-G_M)\beta^{j}, && j\geq M; \end{aligned} \tag{2.8a}$$

$$u_j=\frac{\sigma(\beta-1)\cos j\theta}{2\gamma[\beta\cos(M-1)\theta-\cos M\theta]}$$

$$+\frac{(\beta-1)\sin j\theta}{2[\sin M\theta-\beta\sin(M-1)\theta]}-\frac{\sigma}{2\gamma},\quad -M<j<M;$$

where

$$F_M=\frac{\cos M\theta-\cos(M-1)\theta}{2\gamma[\beta^M\cos(M-1)\theta-\beta^{M-1}\cos M\theta]},$$

$$G_M=\frac{\sin M\theta-\sin(M-1)\theta}{2[\beta^M\sin(M-1)\theta-\beta^{M-1}\sin M\theta]}. \tag{2.8b}$$

In order to complete the solution for the displacements, it is necessary to determine the value of M, characterizing the number of weak bonds, which corresponds to given values of P, the force constant ratio, $\gamma=\phi/b$, and the applied stress, σ. We proceed in an inverse fashion, regarding M as well as P as prescribed, and determine the range of γ for which a solution with $n_w=2M-1$ weak bonds is (1) compatible with the assumed force law (2.2b) and (2) is mechanically stable.

For compatibility, the following conditions are necessary [see Eq. (2.2b)]:

$$\begin{aligned} &\bar{v}_M\leq\phi, && \bar{v}_{-M}\geq-\phi,\\ &\bar{v}_{M-1}\geq-\left(\frac{b}{2}-\phi\right), && \bar{v}_{-M+1}\leq\left(\frac{b}{2}-\phi\right); \end{aligned} \tag{2.9a}$$

these will later be shown to lead to sufficient conditions. With the use of Eqs. (2.4) and (2.8) these inequalities become

$$\gamma\geq-\beta^MG_M+\sigma(1+\beta^MF_M)\,,\gamma\geq-\beta^MG_M-\sigma(1+\beta^MF_M);$$

and (2.9b)

$$\begin{aligned} &\gamma\leq-\beta^{M-1}G_M+\sigma(1+\beta^{M-1}F_M)\,,\\ &\gamma\leq-\beta^{M-1}G_M-\sigma(1+\beta^{M-1}F_M)\,. \end{aligned}$$

Clearly, if a given bond configuration is not compatible without stress it is not compatible under stress. The stress-free compatibility range is

$$-\beta^MG_M\leq\gamma\leq-\beta^{M-1}G_M\,. \tag{2.9c}$$

Through the use of Eqs. (2.7) and (2.8b) together with some trigonometric manipulation, the following implicit

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

equation for the lower limit on γ is obtained:

$$\tan M\theta=\frac{1-\beta}{1+\beta}\operatorname{ctn}\frac{\theta}{2}. \tag{2.9d}$$

Let the value of γ^{-1} that satisfies this equation be denoted by C_M. The corresponding equation for the upper limit on γ is

$$\tan(M-1)\theta=\frac{1-\beta}{1+\beta}\operatorname{ctn}\frac{\theta}{2} \tag{2.9e}$$

so that the range corresponding to that of Eq. (2.9c) is

$$C_{M-1}\leq\gamma^{-1}\leq C_M. \tag{2.9f}$$

A detailed investigation shows that some of the inequalities (2.2b) other than those of (2.9a) will be violated unless the lowest values of θ satisfying Eqs. (2.9d) and (2.9e) are chosen. If this is done, then the maximum value of θ in the range (2.9f) satisfies the inequality $(M-1)\theta<\pi/2$, and this can be used together with Eqs. (2.8) to show that for $0<j<M$ u_j is negative and $u_{j+1}<u_j$ and that for $j\geq M$ u_j is positive and $u_{j+1}<u_j$ (only positive j, because of symmetry, need be considered). Then all of the inequalities (2.2) are satisfied.

For the examination of the stability of the equilibrium configuration defined by the displacements u_j it is convenient to return first to the consideration of a finite chain, $-(N+1)\leq j\leq N+1$ [where N may be a half-odd-integer; see Fig. 1(e)] and then to let $N\rightarrow\infty$. Let q_j be a small displacement of the jth atom from the equilibrium position being examined, with $q_{-N-1}=q_{N+1}=0$, $V(q_{-N},\cdots,q_N)$ the resulting change in potential energy, and $S_{ij}=(1/k_1)(\partial^2V/\partial q_i\partial q_j)$, $-N\leq i,\ j\leq N$, evaluated at $q_j\equiv 0$. Then

$$\begin{aligned}
&S_{jj}=(2+P);\quad -N\leq j\leq -M;\quad M\leq j\leq N;\\
&S_{jj}=(2-Q);\quad -M<j<M;\\
&S_{j,j+1}=S_{j+1,j}=-1;\quad -N+1\leq j\leq N-1;\\
&S_{ij}=0\quad\text{for}\quad |i-j|>1.
\end{aligned} \tag{2.10}$$

Let $q(\lambda)=(q_{-N}(\lambda),q_{-N+1}(\lambda),\cdots,q_N(\lambda))$ be an eigenvector of S_{ij} corresponding to the eigenvalue of λ. Then $q_j(\lambda)$ are the solutions to the difference equations

$$\begin{aligned}
&q_{j-1}-(2+P-\lambda)q_j+q_{j+1}=0;\\
&\qquad -N\leq j\leq -M;\quad M\leq j\leq N;\\
&q_{j-1}-(2-Q-\lambda)q_j+q_{j+1}=0;\\
&\qquad -M<j<M;\quad q_{-N-1}=q_{N+1}=0.
\end{aligned} \tag{2.11}$$

In the terminology of Gantmacher[7] S_{ij} is a normal Jacobi matrix. The eigenvalues λ are then real and simple and are denoted by $\lambda_0<\lambda_1<\cdots\lambda_{2N+1}$. A second property of normal Jacobi matrices is that the sequence $q_{-N}(\lambda_p)$, $q_{-N+1}(\lambda_p)$, $\cdots$, $q_N(\lambda_p)$, $p=0,\cdots,2N+1$, has precisely p changes in sign. In order to determine criteria for the positive definiteness of S_{ij} and hence for the stability of a solution with $2M-1$ weak bonds, it is sufficient to consider λ_0 and we therefore seek a solution to Eqs. (2.11) with no changes in sign. This solution, for the case $N\rightarrow\infty$ may be written in the form, analogous to that of Eq. (2.6),

$$q_j=A\bar\beta^{-j},\quad j\leq -M+1; \tag{2.12a}$$

$$q_j=B\bar\beta^{j},\quad j\geq M-1;$$

$$q_j=C\cos j\bar\theta+D\sin j\bar\theta,\quad -M\leq j\leq M; \tag{2.12b}$$

where

$$\begin{aligned}
&\bar\beta=\tfrac{1}{2}[\bar P+2-(\bar P^2+4\bar P)^{1/2}];\\
&\cos\bar\theta=1-\bar Q/2;\\
&\bar P=P-\lambda,\quad \bar Q=Q+\lambda.
\end{aligned} \tag{2.13}$$

The four coefficients, A, B, C, and D are determined by the requirement that Eqs. (2.12a) and (2.12b) agree for the four atoms for which the expressions overlap. The resulting system of four homogeneous equations may be separated into two sets of two each for symmetric and antisymmetric eigenvectors, respectively. We consider only the former set of equations since λ_0 corresponds to a symmetric eigenvector (as noted above, there are no changes in sign in the sequence of its components). They are [with $F=\frac{1}{2}(A+B)$]

$$\begin{aligned}
&C\cos M\bar\theta-F\bar\beta^M=0,\\
&C\cos(M-1)\bar\theta-F\bar\beta^{M-1}=0,
\end{aligned} \tag{2.14}$$

and equating the determinant of the system to zero leads to the following implicit equation for λ_0,

$$\bar\beta=\frac{\cos M\bar\theta}{\cos(M-1)\bar\theta}, \tag{2.15a}$$

where λ_0 is that root for which $M\bar\theta<(\pi/2)$. We consider next the variation of $\lambda_0=\lambda_0(\gamma^{-1};M)$ with γ^{-1}, with P constant. The following sequence of inequalities may be determined from the equations cited:

$$\frac{d\bar\beta}{d\lambda_0}>0, \qquad \text{[Eq. (2.13)]}$$

$$\frac{d\bar\beta}{d(\cos\bar\theta)}>0. \qquad \text{[Eq. (2.15a)]}$$

Consequently $d(\cos\bar\theta)/d\lambda_0>0$, and from Eqs. (2.4a) and (2.13)

$$\frac{d\gamma^{-1}}{d\lambda_0}=\frac{d\gamma^{-1}}{dQ}\frac{dQ}{d\lambda_0}=-\frac{2P}{Q^2}\left(\frac{dQ}{d\lambda_0}-\frac{1}{2}\right)=\frac{2P}{Q^2}\left[\frac{d(\cos\bar\theta)}{d\lambda_0}+\frac{1}{2}\right]>0.$$

The same result, that a decrease in γ^{-1} decreases λ_0,

[7] F. R. Gantmacher and M. G. Krein, *Oszillationsmatrizen, Oszillationskerne und Kleine Schwingungen Mechanischer Systeme* (Akademie-Verlag, Berlin, 1960), p. 80.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

may also be seen directly from Rayleigh's theorems,[8] since, as appears from Eqs. (2.4a), a decrease in γ^{-1} corresponds to an algebraic decrease of the weak-bond spring constants. Therefore the solution corresponding to $2M-1$ weak bonds will be stable only for $\gamma^{-1}\geq S_M$ where S_M is that value of γ^{-1} for which $\lambda_0(S_M;M)=0$. Since for $\lambda_0=0$, $\bar{\beta}=\beta$ and $\bar{\theta}=\theta$, the transcendental equation for S_M is

$$\cos M\theta-\beta\cos(M-1)\theta=0, \quad M\theta<(\pi/2); \quad (2.15b)$$

or

$$\tan(M-\tfrac{1}{2})\theta=\frac{1-\beta}{1+\beta}\operatorname{ctn}\frac{\theta}{2}, \quad (2.15c)$$

where Eq. (2.15c) follows upon use of trigonometric identities. Consequently $S_M=C_{M-\frac{1}{2}}$.

Therefore, combining the discussions for compatibility and stability, we may conclude that for a value of γ^{-1} in the range $C_{M-\frac{1}{2}}\leq\gamma^{-1}\leq C_M$ a solution with $2M-1$ weak bonds is both compatible and stable, while the only other compatible solution for the same γ, that with one more weak bond, is unstable.

Peierls Stress

Consider next a linear chain with displacements corresponding to a compatible and stable solution with $2M-1$ weak bonds and apply to it a stress σ which is zero initially and increases slowly (so that inertia effects are negligible). Equations (2.9b) show that this bond configuration will remain compatible while

$$\sigma\leq\frac{\gamma+\beta^M G_M}{1+\beta^M F_M}, \quad \sigma\leq-\frac{\gamma+\beta^{M-1}G_M}{1+\beta^{M-1}F_M}. \quad (2.16a)$$

We define σ_P as the stress above which these inequalities are invalid; for $\sigma>\sigma_P$ the problem has no (statically) compatible solution, i.e., the chain could not support a stress $\sigma>\sigma_P$ in static equilibrium. Therefore σ_P corresponds to the Peierls stress for the model considered. The first of Eqs. (2.16a) may be shown to be controlling; this yields

$$\sigma_P=\gamma\left[\frac{\cos M\theta-\beta\cos(M-1)\theta}{\sin M\theta-\beta\sin(M-1)\theta}\right]\times\left[\frac{\sin M\theta-\delta\sin(M-1)\theta}{\cos M\theta-\delta\cos(M-1)\theta}\right], \quad (2.16b)$$

where $\delta=\beta(2\gamma-1)/(2\gamma-\beta)$. We note that $\sigma_P=0$ for $\gamma^{-1}=C_M$ [this follows from Eq. (2.9d)] and for $\gamma^{-1}=S_M$ [Eq. (2.15b)], that is at the limits of compatibility and stability of configurations with $2M-1$ weak bonds.

Numerical Results

Equation (2.16b) has been evaluated, for values of P (the ratio of noncentral to central forces) equal to 1, 0.5, and 0.2. For each value of P the parameter γ (the ratio of perfect crystal shear strength to shear modulus) was varied over a range wide enough to show the important details of the problem. The results are shown in Figs. 2 and 3; for purposes of comparison Sanders' results,[4] corresponding to $P=0.5$, are also shown in Fig. 2.

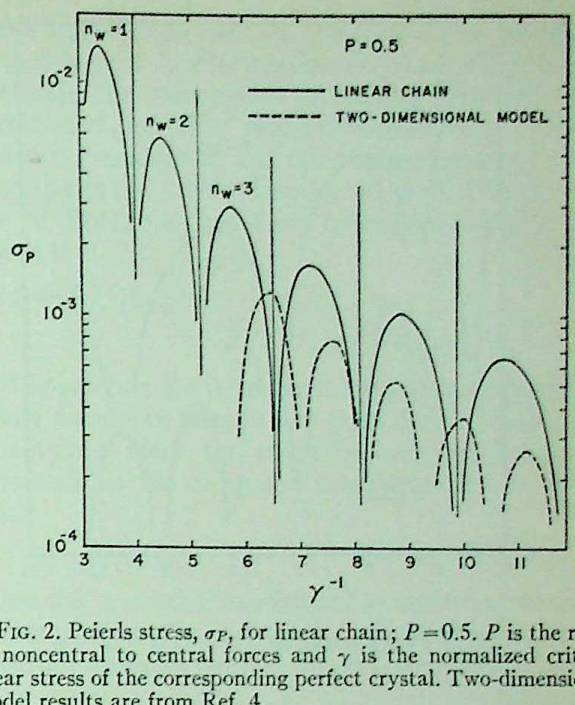


FIG. 2. Peierls stress, σ_P, for linear chain; $P=0.5$. P is the ratio of noncentral to central forces and γ is the normalized critical shear stress of the corresponding perfect crystal. Two-dimensional model results are from Ref. 4.

The agreement between these results and the earlier ones is quite good, in view of the expected sensitivity of the Peierls stress to the details of the model. Both sets have the quasicyclic pattern, and both have about the same exponential dependence on γ^{-1}, while the present results have peak values roughly twice those of the earlier ones.

A qualitative argument would indicate two opposite effects in passing from the one to the two-dimensional model: if the single chain treated here is considered as the first atom row above the slip plane, then the presence of the rows below it would tend to narrow the dislocation, increasing the Peierls stress, while the rows above it would increase the width and lower the stress. Apparently, the latter effect dominates.

The agreement between the one- and two-dimensional cases might serve to substantiate the superiority of the accurate calculation on a discrete model, even a relatively poor model, over the use of a continuum model for dislocation studies.[9]

3. STATISTICAL MECHANICAL CONSIDERATIONS

We wish to consider next the creep, due to thermal motion, of a linear chain, as described in Sec. 2, with a dislocation under an applied stress $\sigma<\sigma_P$. For this

[8] Lord Rayleigh, *The Theory of Sound* (Dover Publications, New York, 1945), Vol. 1, p. 111 *et seq.*

[9] See also R. Hobart and V. Celli, J. Appl. Phys. 33, 60 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

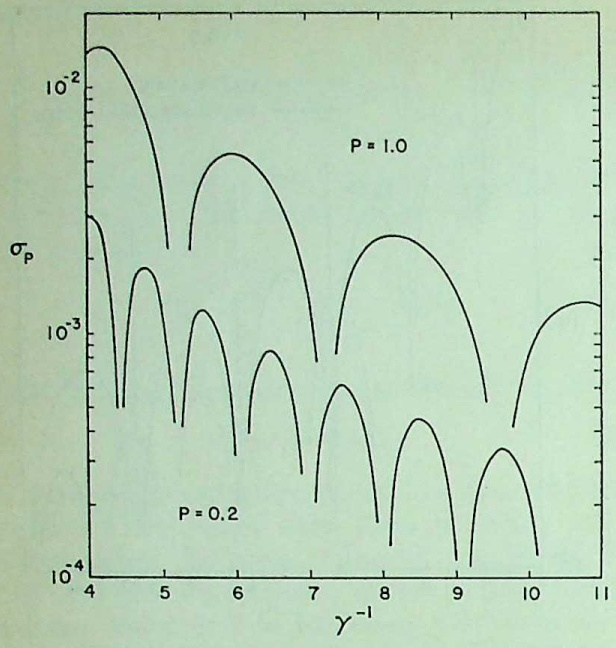


FIG. 3. Peierls stress, σ_P, for $P=1.0$, 0.2.

purpose an ensemble of such chains is assumed with distribution function $\rho(p_{-N},\cdots,p_N,q_{-N},\cdots,q_N)=\rho(p,q)$ where $q_j(t)$ is the displacement at time t of the jth atom from its equilibrium position under stress σ, and $p_j=m\dot{q}_j$ is the momentum of the jth particle with mass m. We treat a chain in thermal equilibrium with its surroundings at temperature T and accordingly set

$$\rho(p,q)=Ce^{-H/kT}, \tag{3.1}$$

where

$$H(p,q)=(1/2m)p_jp_j+V(q_{-N},\cdots,q_N), \tag{3.2}$$

$$C^{-1}=\int_\Omega e^{-H(p,q)/kT}\prod_i dp_i\,dq_i. \tag{3.3}$$

In these equations and in the remainder of this section latin subscripts assume the values $-N,\cdots,N$, where by the latter notation we signify zero included. Greek subscripts assume the values $-N,\cdots,-1,1,\cdots,N$, zero excluded. The symbol Ω denotes the full phase space, $-\infty<p_i<\infty$, $-\infty<q_i<\infty$, and the summation convention is employed.

The potential energy function $V(q)$ for the type of system under consideration admits of a number of positions of stable equilibrium with intermediate positions of unstable equilibrium. Let a typical stable equilibrium position have coordinates $q=(q_{-N},\cdots,q_N)=0$. In a neighborhood of $q=0$, the potential $V(q)$ may be approximated by

$$V(q)=(k_1/2)S_{ij}q_iq_j, \tag{3.4}$$

where the matrix $S=[S_{ij}]$, $S_{ij}=(1/k_1)(\partial^2V(0)/\partial q_i\partial q_j)$, is positive-definite. It is now assumed that the system spends most of the time in this neighborhood of $q=0$, so that $\rho(p,q)$ is very small elsewhere, and that therefore for the purpose of computing the normalization constant C, Eq. (3.3), the approximate expression for $V(q)$, Eq. (3.4), may be used. The result is

$$C=\frac{|S|^{1/2}k_1^{N+\frac{1}{2}}}{(2\pi kT)^{2N+1}m^{N+\frac{1}{2}}}, \tag{3.5}$$

where $|S|$ is the determinant of the matrix S.

Consider next the position of unstable equilibrium at $q=d=(d_{-N},\cdots,d_N)$ nearest to the stable equilibrium position at $q=0$. We confine attention to a neighborhood of the saddle point $q=d$ in which the potential $V(q)$ is adequately represented by its quadratic approximation

$$V(q)=V(d)+(k_1/2)U_{ij}(q_i-d_i)(q_j-d_j), \tag{3.6}$$

where

$$U_{ij}=(1/k_1)(\partial^2V(d)/\partial q_i\partial q_j) \tag{3.6a}$$

so that the equations of motion in that neighborhood take the form

$$m\ddot{q}_i=-k_1U_{ij}(q_j-d_j). \tag{3.7}$$

Let λ_i be the eigenvalues of the matrix $U=[U_{ij}]$ with associated unit eigenvectors a_{ij} and let $\lambda_0<0$ be the minimum eigenvalue with the sign of a_{0j} chosen so that $a_{0j}d_j>0$. It is assumed that λ_0 is the only negative eigenvalue.[10] With the orthogonal change in coordinates

$$Q_i=a_{ij}(q_j-d_j), \tag{3.8}$$

Eq. (3.7) assumes the form, for $Q_0(t)$,

$$\ddot{Q}_0=+s^2Q_0,\quad s^2=-k_1\lambda_0/m, \tag{3.9}$$

with the solution

$$Q_0(t)=\tfrac{1}{2}[Q_0(0)+s^{-1}\dot{Q}_0(0)]e^{st}+\tfrac{1}{2}[Q_0(0)-s^{-1}\dot{Q}_0(0)]e^{-st}. \tag{3.10}$$

Therefore, if at an arbitrary instant of time, $t=0$,

$$Q_0(0)=0\quad\text{and}\quad\dot{Q}_0(0)>0,$$

then $Q_0(t)$ will increase monotonically, at least as long as Eq. (3.5) applies. Since a_{0j} points away from the stable equilibrium position at $q=0$, this result implies that if the system, in its notion in phase space, reaches the hyperplane $Q_0=0$ with positive velocity, it will not return to its original equilibrium position but continue on towards the next equilibrium position.[11] This hyperplane corresponds therefore to the crest of the ridge between the two valleys. In order to compute the frequency, f, with which the system reaches the hyperplane $Q_0=0$ with $\dot{Q}_0>0$, we introduce, as new coordinates for phase space the Q_i as defined by Eq. (3.8) and the corresponding new momenta $P_i=a_{ij}p_j$. The distribution function in terms of these new variables will be denoted by $\rho(P_0,Q_0,P_\alpha,Q_\alpha)$ where the arguments are separated as shown for future convenience. The

[10] For the case of the linear chain described in Sec. 2, it can be proved that only one negative eigenvalue exists.

[11] This statement is, of course, intended only to apply to short times.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

frequency f may now be obtained directly by use of a formula derived by Rice.[12] This formula, which requires the assumption of the ergodic hypothesis for its validity, is in the present case,

$$f=\frac{1}{m}\int_0^\infty P_0 dP_0 \int_{\Omega'} \rho(P_0,0,P_\alpha,Q_\alpha)\prod_\alpha dP_\alpha dQ_\alpha, \quad (3.11)$$

where Ω' is the subspace, $-\infty<P_\alpha<\infty$, $-\infty<Q_\alpha<\infty$. This leads, after the integrations in P_i are performed, to

$$f=\frac{k_1{}^{N+\frac{1}{2}}|S|^{1/2}}{(2\pi)^{N+1}(kT)^N\sqrt{m}} \times\int_{\Omega''}\exp\left\{\frac{-V(0,Q_\alpha)}{kT}\right\}\prod_\alpha dQ_\alpha, \quad (3.12)$$

where Ω'' is the subspace $-\infty<Q_\alpha<\infty$ and $V(Q_0,Q_\alpha)$ is the function $V(q)$ expressed in terms of the new coordinates Q_i. The assumption is next made that, for the purpose of evaluating the integral in Eq. (3.12), $V(0,Q_\alpha)$ may be evaluated from Eq. (3.6), the expression valid in the neighborhood of the saddle point. Then

$$V(0,Q_\alpha)=V(d)+(k_1/2)\textstyle\sum_\alpha \lambda_\alpha Q_\alpha{}^2. \quad (3.13)$$

Substitution of this expression into Eq. (3.12) leads to the following expression for f:

$$f=f_0\nu e^{-V(d)/kT} \quad (3.14)$$

where

$$f_0=\frac{1}{2\pi}(k_2/m)^{1/2}; \quad \nu=\left(\frac{\lambda_0|S|}{P|U|}\right)^{1/2} \quad (3.14a)$$

and where $|U|=\prod_i\lambda_i$ is the determinant of the matrix U. The quotient of determinants may be written

$$|S||U|^{-1}=|I+S^{-1}D|^{-1}, \quad (3.15)$$

where S^{-1} is the matrix inverse to S, I is the unit matrix and $D=U-S$. In the type of problem considered here the difference matrix, D, will have only a few nonzero elements; in this case, the determinant $|I+S^{-1}D|$ may be expressed in terms of a correspondingly low-order determinant, as will be seen in Sec. 4.

4. CREEP OF LINEAR CHAIN

In this section, the rate formula of Eq. (3.14) is applied to the linear chain described in Sec. 2 under an applied stress $\sigma<\sigma_P$.

Saddle-Point Energy, $V(d)$

In Sec. 3, two approximate expressions for $V(q)$, Eqs. (3.4) and (3.6) were given, the first valid in the neighborhood of the stable equilibrium position, $q=0$, and the second valid in the neighborhood of the unstable equilibrium position, $q=d$. For the model described in Sec. 2, with a piecewise quadratic substrate potential, the two regions of validity (referred to as stable and unstable regions, respectively) have a common boundary B, and the displacements $q_i{}^P$ corresponding to the Peierls stress σ_P fall on B. Therefore, by use of Eqs. (3.4) and (3.6) to evaluate $V(q^P)$, it is found that

$$V(d)=(k_1/2)S_{ij}q_i{}^Pq_j{}^P -(k_1/2)U_{ij}(q_i{}^P-d_i)(q_j{}^P-d_j). \quad (4.1)$$

The position $q_i=0$ corresponds to stable equilibrium under an applied stress $\bar\sigma<\bar\sigma_P$ (bars denote dimensional quantities). Since the stress-displacement relation is linear within the stable and unstable region, it follows that

$$\partial V(q^P)/\partial q_i=k_1S_{ij}q_j{}^P=\bar\sigma_P-\bar\sigma, \quad \text{all } i. \quad (4.2)$$

when the position q^P is regarded as arrived at from the stable position. Similarly, if it is regarded as arrived at from the unstable position at $q=d$, we have

$$\partial V(q^P)/\partial q_i=k_1U_{ij}(q_j{}^P-d_j)=\bar\sigma_P-\bar\sigma, \quad \text{all } i. \quad (4.3)$$

Substitution from Eqs. (4.2) and (4.3) into Eq. (4.1) yields

$$V(d)=\tfrac{1}{2}(\bar\sigma_P-\bar\sigma)\textstyle\sum_i d_i. \quad (4.4)$$

Because of the linear behavior in both regions, $q_i{}^P(\bar\sigma)$, the displacements to the common B from the equilibrium positions corresponding to the applied stress, $\bar\sigma$, are given by

$$q_i{}^P(\bar\sigma)=[(\bar\sigma_P-\bar\sigma)/\bar\sigma_P]q_i{}^P(0),$$
$$d_i(\bar\sigma)-q_i{}^P(\bar\sigma)=[(\bar\sigma_P-\bar\sigma)/\bar\sigma_P][d_i(0)-q_i{}^P(0)],$$

so that

$$d_i(\bar\sigma)[(\bar\sigma_P-\bar\sigma)/\bar\sigma_P]d_i(0). \quad (4.5)$$

Therefore, the energy at the saddle point $d(\bar\sigma)$, corresponding to an applied stress, $\bar\sigma$, is

$$V(d(\bar\sigma))=[(\bar\sigma_P-\bar\sigma)^2/2\bar\sigma_P]\textstyle\sum_i d_i(0). \quad (4.6)$$

Consider $V(d(0))$. Let Δ_j be the displacement of the jth atom when the dislocation moves from the stable equilibrium position at $q=0$ to the next adjacent stable equilibrium position. Then $V(0)=V(\Delta)=0$, and $V(d(0))$ can be computed by considering either displacements from $q=0$ or $q=\Delta$. In both cases Eq. (4.6) applies (with oppositely signed $\bar\sigma_P$) so that

$$V(d(0))=\tfrac{1}{2}\bar\sigma_P\textstyle\sum_i d_i(0)=-\tfrac{1}{2}\bar\sigma_P\sum_i(d_i(0)-\Delta_i)$$

or

$$2\textstyle\sum_i d_i(0)=\sum_i\Delta_i=b, \quad (4.7)$$

where b is the Burger's vector of the dislocation. Substitution into Eq. (4.6) then yields the following formula for the saddle-point energy:

$$v(d(\bar\sigma))=\frac{(\bar\sigma_P-\bar\sigma)^2b}{4\bar\sigma_P}=\frac{(\sigma_P-\sigma)^2}{4\sigma_P}k_2b^2. \quad (4.8)$$

[12] S. O. Rice, Bell Syst. Tech. J. 24, 58 (1945). Reprinted in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover Publications, New York, 1954), p. 190.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

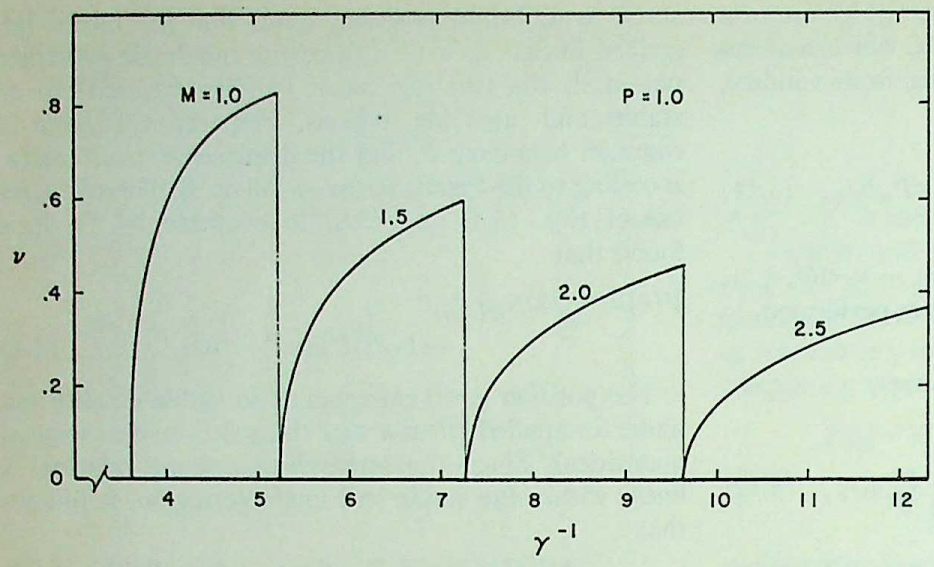


FIG. 4. Variation of the dimensionless frequency ν, Eq. (3.14), with the normalized critical shear stress γ.

Computation of ν

For the model of Sec. 2 which is stable with $2M-1$ weak bonds, only the elements of S and U corresponding to $i=j=M$ differ, so that the only nonzero element of $D=U-S$ is

$$D_{MM}=-(P+Q)=\frac{-P}{1-2\gamma}. \tag{4.9}$$

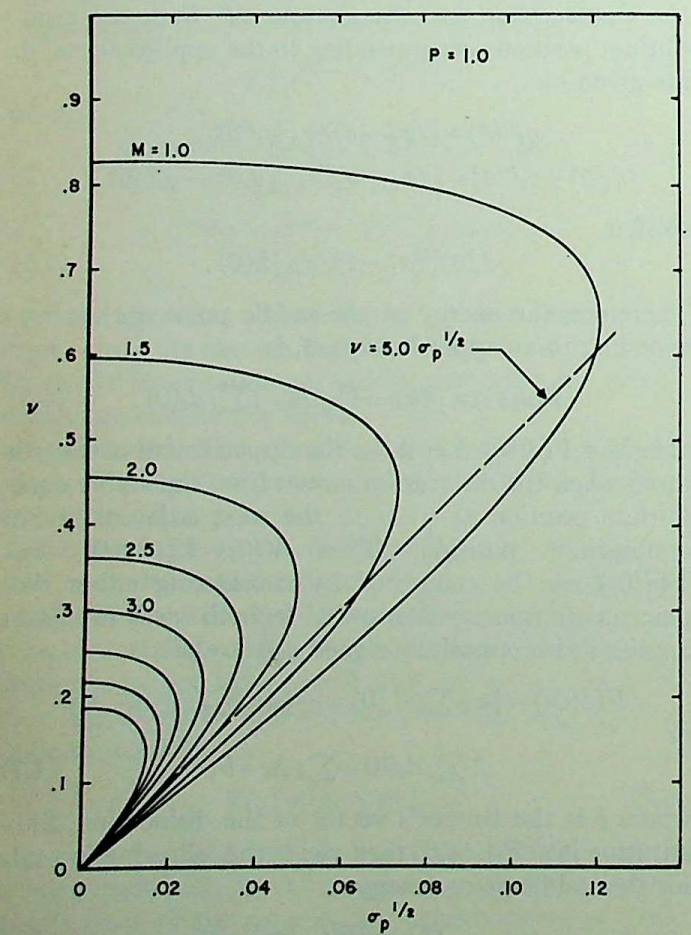


FIG. 5. Variation of the dimensionless frequency ν, Eq. (3.14), with Peierls stress, σ_P.

Substitution in Eq. (3.15) gives the result

$$|S||U|^{-1}=(1+S^{-1}{}_{MM}D_{MM})^{-1} \tag{4.10}$$

so that only the element $S^{-1}{}_{MM}$ of the matrix S^{-1} is needed. This is found (for the infinite chain) by solution of the difference equations $S_{ij}w_j=\delta_{iM}$ with the result

$$S^{-1}{}_{MM}=\frac{\beta}{2}\left[\frac{\cos M\theta}{\cos M\theta-\beta\cos(M-1)\theta}+\frac{\sin M\theta}{\sin M\theta-\beta\sin(M-1)\theta}\right]. \tag{4.11}$$

Finally it is necessary to calculate λ_0, the minimum eigenvalue of the matrix U. This is found from Eq. (2.15a) with M replaced by $M+\frac{1}{2}$. For the case in which stability corresponds to $M=1$, an explicit formula for λ_0 may be derived, namely,

$$\lambda_0=\left[\frac{(2-Q)(P+Q+2)-P-4}{P+Q+1}\right]. \tag{4.12}$$

For other values of M, numerical evaluation of Eq. (2.15a) was used. The resulting values of ν, the dimensionless frequency that appears in the rate equation, Eq. (3.14), are plotted in Fig. 4 for $P=1.0$ as a function of γ. It may be noted that $\nu=0$ for those values of γ which correspond to the limit of stability, S_M, for the given configuration, as does σ_P (Figs. 2 and 3). However, ν increases monotonically within the range $S_M\leq\gamma^{-1}<C_M$, whereas σ_P returns to zero at $\gamma^{-1}=C_M$. It is clear, therefore, that these calculations do not reveal a linear dependence of ν upon $\sqrt{\sigma_P}$ as might be expected from a model treating the dislocation as a single entity moving in a potential field corresponding to σ_P. A plot of ν as a function of $\sqrt{\sigma_P}$, given in Fig. 5, reveals, however, that if only the initial portion of each range of γ is considered, the two quantities are indeed linearly related. It may be that these deviations from

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Comparison of theoretical and analog creep rates.

k_2b^2/kT		990	560	245
f/f_0	analog	0.157	0.22	0.46
	theoretical	0.456	0.53	0.595

linearity are due in part to the discontinuous substrate potential employed. If only the initial portions of each range of γ are considered, then the present calculations lead to the relationship $v \simeq a\sqrt{\sigma_P}$ with $a=5\pm0.5$ for $0.1 \leq P \leq 10$. This may be compared with the similar relationships found, for example, by Seeger,[13] Weertman,[14] and Lothe[1] in their treatments of a dislocation as a vibrating string, the latter two obtaining values of $a=5.6$ and 2.5 respectively.

Analog Computer Results

A REAC Model-400 analog computer was used in an attempt to check these rate calculations, at least in order of magnitude. Because of limitations in equipment only a six-atom model could be simulated; also it was not considered feasible to run through a random sequence of initial conditions, so initial conditions were chosen that would give a minimum frequency of jumps of the dislocation. This model, then, was run for $P=1$, $\gamma=0.2396$, and $\sigma=0.010$ ($\sigma_P=0.0146$), and for three values of temperature. Table I shows these results together with corresponding results based on Eq. (3.14).

The analog data do agree, within a factor of 3, with the theoretical results. The analog results show a lower frequency, as expected (because of initial conditions). They also exhibit a more sensitive dependence on temperature; this again was expected, because of the fact that only six atoms were simulated. In view of the limitations on the analog model, the agreement seems to be good, and appears to lend justification to the assumptions made in the theoretical derivation.

[13] A. Seeger, Phil. Mag. 1, 651 (1956).
[14] J. Weertman, J. Appl. Phys. 28, 1185 (1957).

5. CONCLUSIONS

Some conclusions which may be drawn from this work are as follows:

(1) The one-dimensional dislocation model yields values for the Peierls stress σ_P in surprisingly good agreement with those for the two-dimensional model considered by Sanders.[4]

(2) In particular, the sensitivity of σ_P to small changes in γ, the critical shear stress, which was found in the two-dimensional calculations[4] is confirmed here.

(3) Detailed results could be obtained on rate of motion of the dislocation under various levels of applied stress less than the Peierls stress and for a range of values of the other parameters of the problem. The relation between these rates and the Peierls stress is in good agreement with that found earlier by treating the dislocation as a vibrating string, except for systematic deviations that are possibly due in part to the discontinuous substrate potential used here.

(4) The rate formula, Eq. (3.14), which has been applied here to a specific one-dimensional model has broader applicability. It allows explicit calculation of rates of transition, without consideration of the concept of entropy of activation, when the potential energy function of the system as a function of the microscopic coordinates is known.

ACKNOWLEDGMENT

The authors wish to thank Mrs. J. H. Clearman who programmed and carried out the numerical calculations.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Incoherent and Thermal Small-Angle X-Ray Scattering in Silicon

D. L. WEINBERG*
Corning Glass Works, Corning, New York
(Received 30 December 1963)

The intensity of small-angle x-ray scattering in single-crystal silicon has been measured with the scattering vector **K** oriented along each of the three symmetry directions. The scattering intensity extrapolates to zero angle in reasonable agreement with the thermal scattering, calculated from the velocity of sound in each direction. The thermal scattering is calculated from known phonon dispersion relations and subtracted from the total to obtain the incoherent, or Compton, scattering. The incoherent scattering is smaller than that calculated by Freeman from Hartree-Fock wave functions for the free atom. This conclusion agrees with previous approximate measurements at small K on other crystals. Theoretical ideas about the source of the deficiency, apparently a solid-state effect, and about small-angle thermal scattering are outlined. The incoherent scattering shows little anisotropy and no sharp rise with increasing angle, which would suggest the onset of excitation of electrons across the forbidden band.

I. INTRODUCTION AND THEORY

A. Incoherent Scattering

IN a perfect, rigid crystal all the diffuse single scattering is incoherent. From the semiclassical or quantum theory, it is[1]

$$I_{\rm inc}(\mathbf{K})=\sum_{m\neq 0}\left|\int\psi_m^*\psi_0\sum_j^N e^{i\mathbf{K}\cdot\mathbf{r}_j}d\tau\right|^2 \quad (1)$$

$$=\int|\psi_0|^2|\sum_j^N e^{i\mathbf{K}\cdot\mathbf{r}_j}|^2 d\tau - \left|\int|\psi_0|^2\sum_j^N e^{i\mathbf{K}\cdot\mathbf{r}_j}d\tau\right|^2 \quad (2)$$

in electron units (e.u.). $K=4\pi\sin\frac{1}{2}\theta/\lambda$, θ is the scattering angle, λ is the x-ray wavelength, ψ_m is the wave function in state m, $m=0$ is the ground state, N is the total number of electrons, $\mathbf{r}_j$ is the position of the jth electron, and the integrations are over all coordinates of the electrons. Then $I_{\rm inc}(0)=0$. The approximations that the x-ray frequency is small enough to avoid recoil and relativity corrections, but that it is much larger than the Bohr frequency ν_{0m} for any m which contributes appreciably to the sum in (1), are used. The former assumption holds in the present work. The latter assumption excludes anomalous dispersions, or Raman, terms which correspond to excited final electronic states. These terms are too small to have been observed in x-ray scattering, but are finite at $K=0$.

Thermal motion produces additional diffuse scattering $I_{\rm th}$. Walker[2] separated $I_{\rm inc}$ from $I_{\rm th}$ in Al experimentally, by their temperature dependences. $I_{\rm th}$ is almost unshifted in wavelength, but $I_{\rm inc}$ is known, from spectral measurements at large K, to peak near a wavelength shift $\Delta\lambda=(h/m_ec)(1-\cos\theta)$, the value for a slow, free electron. Thus Curien and Deroche[3] separated $I_{\rm inc}$ from $I_{\rm th}$ in Al with a dispersive filter. Laval[4] measured $I_{\rm inc}$ in Al, diamond, and KCl by subtracting estimated $I_{\rm th}$ from the total diffuse scattering. Measured phonon dispersion relations were not then available. Cribier[5] treated CaF_2 similarly. Measurements on pure, perfect crystals at small K have been difficult in the past. The weak scattering power requires an unusually strong x-ray source and low background. For convenient λ, θ is small, so that common diffraction equipment is not suitable.

If ψ_0 is a determinant of one-electron functions ψ_j, then from (2),

$$I_{\rm inc}=N-\sum_{jj'}^N\left|\int\psi_j^*(r)\psi_{j'}(r)e^{i\mathbf{K}\cdot\mathbf{r}}d\tau\right|^2. \quad (3)$$

The most careful calculations have used (3) with Hartree-Fock ψ_j for free atoms or ions. A recent review[6] lists the results. They lie above the measurements[4] at small K, but agree otherwise.[6,7] Coherent scattering in crystals (at large K) is also well described by free-atom orbitals.[8] Neglect of interatomic overlap in (3), which is emphasized at small K, presumably produces much of the overestimate of $I_{\rm inc}$. It may be that overlap contributions vanish when $e^{i\mathbf{K}\cdot\mathbf{r}}$ varies with large K, and that well within the atom, ψ_j is similar in free atom and crystal. Overlap effects might be estimated with simple approximate atomic orbitals. The solutions of the Hartree-Fock problem are Bloch functions, extend-

---

* Now at Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts.

[1] R. W. James, *The Optical Principles of the Diffraction of X-Rays* (G. Bell and Sons, Ltd., London, 1948).

[2] C. B. Walker, Phys. Rev. **103**, 558 (1956).

[3] H. Curien and C. Deroche, Bull. Soc. Franç. Minéral. **79**, 102 (1956).

[4] J. Laval, Bull. Soc. Franç. Minéral. **62**, 137 (1939); Compt. Rend. **215**, 278 (1942).

[5] D. Cribier, Rev. Mod. Phys. **30**, 228 (1958); Ann. Phys. **4**, 333 (1959).

[6] H. Curien, in *International Tables for X-Ray Crystallography*, edited by C. H. Macgillavry, G. D. Rieck, and K. Lonsdale (The Kynoch Press, Birmingham, England, 1962), Vol. III, p. 247.

[7] A. J. Freeman, Phys. Rev. **113**, 176 (1959); Acta Cryst. **13**, 190 (1960).

[8] B. W. Batterman, D. R. Chipman, and J. J. DeMarco, Phys. Rev. **122**, 68 (1961); M. Kuriyama and S. Hosoya, J. Phys. Soc. Japan **18**, 1315 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ing over the whole crystal, but the determinental ψ_0 is not changed if one takes each ψ_j as a linear combination of the Bloch functions which is localized on an atom. However, for metallic or covalent bonding, these ψ_j have greater overlap than the free atom orbitals. For example, the assumption of some free electrons,[2] instead of isolated atoms, reduces $I_{\rm inc}$ at small K (except very near $K=0$). In the absence of accurate measurements on crystals at small K, no realistic calculations of $I_{\rm inc}$, including interatomic effects, have been made.

To be Compton scattered in an insulator or semiconductor at absolute zero of temperature, a photon must give up at least the band-gap energy. Suppression of scattering with small $\Delta\lambda$ is reported[9] from spectral measurements on diamond, with a 7-eV band gap. Analogous reduction of $I_{\rm inc}$ at angles less than a minimum, θ_c, calculated from the slow, free-electron Compton relation, has been suggested[2,7,10] as an alternative explanation of the deficiency in $I_{\rm inc}$. Si has a 1.08-eV band gap, and with $\lambda=1.542$ Å, as in this work, $\theta_c=7°$. But the electrons in band-gap solids are not free and slow, so that large energy transfers at small θ may occur. In some cases the deficiency is by comparison with $I_{\rm inc}$ calculated for free closed-shell ions, which themselves have few low-lying excited levels. The idea is inapplicable, *a fortiori*, to a metal.

There is more justification for requiring conservation of quasimomentum, $\mathbf{k}'-\mathbf{k}=\mathbf{K}+\mathbf{K}_L$, from (1). $\mathbf{k}$, $\mathbf{k}'$, are the wave vectors of the excited electron in the initial and final states. $\mathbf{K}_L$ is a vector of the reciprocal lattice. Knowledge of the band structure of Si[11] shows that the addition of this requirement scarcely increases the predicted θ_c, for $\mathbf{K}$ along the [001] direction. However, this kind of description of the excited states seems inadequate.

B. Thermal Scattering

Thermal scattering involves emission or absorption of phonons, but it can be treated classically, and is well understood for harmonically vibrating crystals.[1,12,13]

Since the two primitive sublattices in Si are identical, their amplitudes must be equal in each normal mode. Then the harmonic, one-phonon scattering is[12]

$$I_{\rm th}(\mathbf{K})=\tfrac{1}{2}m^{-1}\exp(-2M)K^2|f|^2 \times\Big[(1+\cos\mathbf{K}_L\cdot\boldsymbol{\varrho})\sum_{i=1}^{3}E_i\omega_i^{-2}\cos(\mathbf{K},\mathbf{e}_i) +(1-\cos\mathbf{K}_L\cdot\boldsymbol{\varrho})\sum_{i=4}^{6}E_i\omega_i^{-2}\cos^2(\mathbf{K},\mathbf{e}_i)\Big] \quad (4)$$

in e.u./atom. The two summations extend over the acoustic and optical modes, respectively. $\mathbf{K}$ is in the zone centered on $\mathbf{K}_L$, $\mathbf{q}=\mathbf{K}-\mathbf{K}_L$, $\omega_i/2\pi$ is the frequency of the ith mode with propogation vector $\mathbf{q}$, $E_i=\frac{1}{2}\hbar\omega_i\times\coth(\hbar\omega_i/2kT)$, $2|f|^2(1+\cos\mathbf{K}\cdot\boldsymbol{\varrho})$ is the usual structure factor, m is the atomic mass, and $(\mathbf{K},\mathbf{e}_i)$ is the angle between $\mathbf{K}$ and the polarization vector $\mathbf{e}_i$. The Debye-Waller factor $\exp(-2M)$ can be neglected in this work, since $2M=0.3\sin^2\theta/\lambda^2$ in Si at 300°K. n-phonon scattering is given by the M^n term of a power series in M. Since M is so small here, multiple-phonon scattering is negligible, and it vanishes in general at $K=0$.

At $K=0$, the optical modes do not contribute, since their phase velocities $\omega_i/K\to\infty$ as $K\to 0$. Nor do purely transverse waves, for which $(\mathbf{q},\mathbf{e}_i)=90°$, contribute when $\mathbf{K}$ is in the first zone, where $\mathbf{q}=\mathbf{K}$. Hence, in directions of such high symmetry that the modes are purely transverse or longitudinal at $\mathbf{q}=0$ (⟨001⟩, ⟨111⟩, and ⟨110⟩ in cubic crystals), the harmonic thermal scattering is

$$I_{\rm th}(0)=m^{-1}|f(0)|^2kTV^{-2} \quad (5)$$

at $K=0$. V is the adiabatic velocity of sound in the direction of $\mathbf{K}$. The correction for anharmonicity is to replace V by the isothermal velocity.[14] The difference is negligible in hard solids.

The calculated thermal scattering in Si is shown in Fig. 1 for $\mathbf{K}$ along each of the three symmetry directions.

II. EXPERIMENTS AND DISCUSSION

A rotating water-cooled Cu anode was used, at 33 kV and 98 mA, with monochromatization by a Ni filter and proportional counter-pulse-height analyzer. Slits 0.6-cm high collimated and analyzed the beam. With the slit widths used for $\theta\geqq 3.5°$, the rms resolution was 1.0° at $\theta=3.5°$, and decreased to 0.7° for $\theta\geqq 10°$. The resulting intense beam was needed to measure the weak scattering of Si at small K. The background was about 4 counts/min, and the smallest sample scattering at $\theta>2°$ was about 100 counts/min. Other experimental aspects have been described.[15] The accuracy of the intensity standardization[15] has been verified[16] within a few percent.

The sample was a (110) disc of 100 Ω-cm n-type Si, made by Philco Corp. Alternate polishing and etching reduced the thickness to 0.0102 cm, uniform to 0.0005 cm over the area used, with x-ray transmission 0.237. The final treatment was an etch, to minimize the damaged surface layer. Static imperfections in this grade of Si should produce completely negligible (volume) diffuse scattering. Low dislocation density was confirmed by Laue photographs and examination of etch pits.

The orientation was accurate to 1°. Since the sample

[9] K. Alexopoulos and G. Brogren, Arkiv Fysik 6, 213 (1953).
[10] J. Laval, Compt. Rend. 215, 359 (1942).
[11] F. Bassani and D. Brust, Phys. Rev. 131, 1524 (1963).
[12] J. Laval, Bull. Soc. Franç. Minéral. 64, 1 (1941).
[13] W. Cochran, Rept. Prog. Phys. 26, 1 (1963); A. A. Maradudin, E. W. Montroll, and G. H. Weiss, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Suppl. 3.
[14] S. M. Rytov, Zh. Eksperim. i Teor. Fiz. 33, 166 (1957) [English transl.: Soviet Phys.—JETP 6, 130 (1958)].
[15] D. L. Weinberg, Rev. Sci. Instr. 34, 691 (1963).
[16] D. L. Weinberg, Phys. Letters 7, 324 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

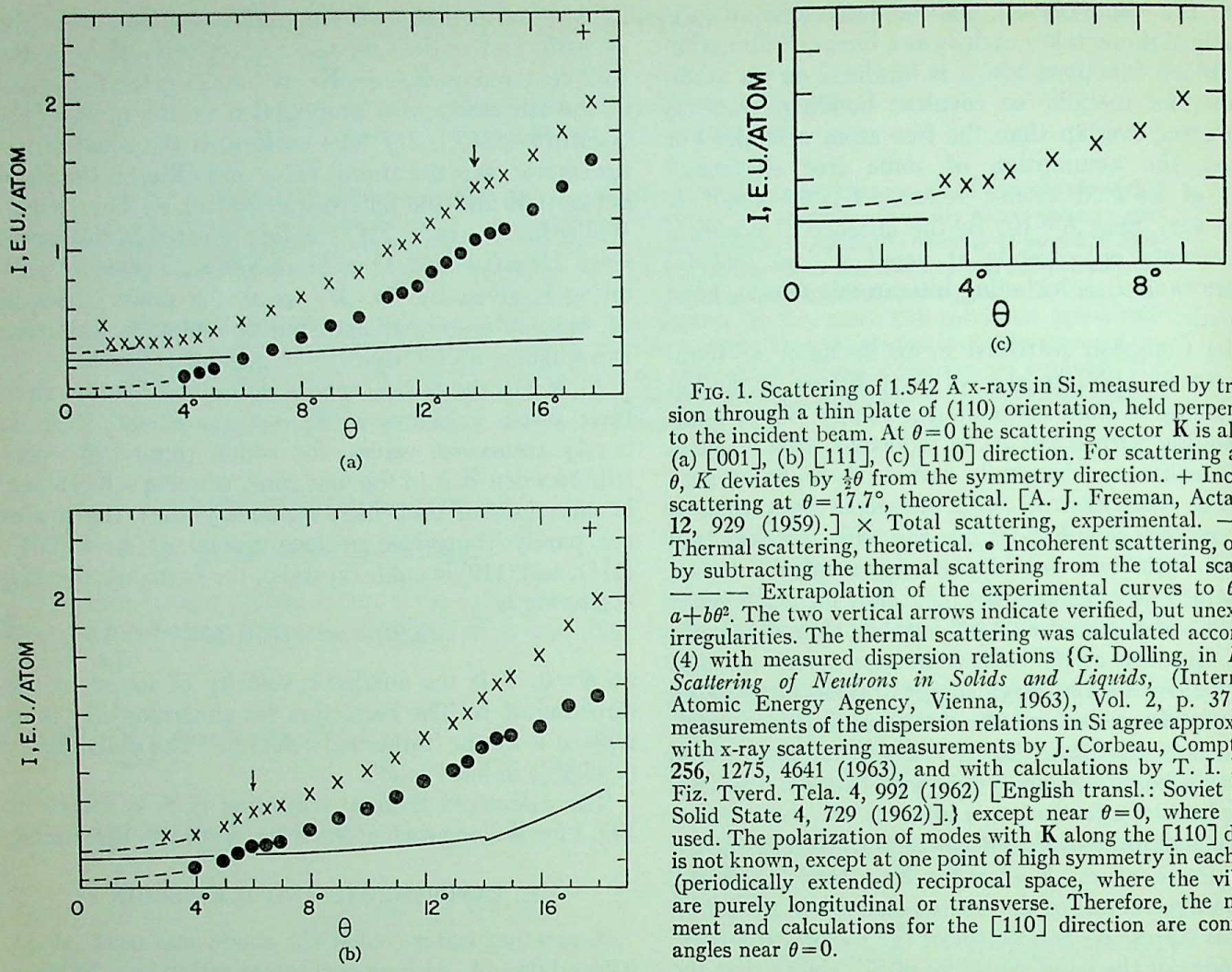


FIG. 1. Scattering of 1.542 Å x-rays in Si, measured by transmission through a thin plate of (110) orientation, held perpendicular to the incident beam. At $\theta=0$ the scattering vector **K** is along the (a) [001], (b) [111], (c) [110] direction. For scattering at angle θ, K deviates by $\frac{1}{2}\theta$ from the symmetry direction. + Incoherent scattering at $\theta=17.7°$, theoretical. [A. J. Freeman, Acta Cryst. 12, 929 (1959).] × Total scattering, experimental. ——— Thermal scattering, theoretical. • Incoherent scattering, obtained by subtracting the thermal scattering from the total scattering. — — — Extrapolation of the experimental curves to $\theta=0$ as $a+b\theta^2$. The two vertical arrows indicate verified, but unexplained irregularities. The thermal scattering was calculated according to (4) with measured dispersion relations {G. Dolling, in *Inelastic Scattering of Neutrons in Solids and Liquids*, (International Atomic Energy Agency, Vienna, 1963), Vol. 2, p. 37. These measurements of the dispersion relations in Si agree approximately with x-ray scattering measurements by J. Corbeau, Compt. Rend. 256, 1275, 4641 (1963), and with calculations by T. I. Kucher, Fiz. Tverd. Tela. 4, 992 (1962) [English transl.: Soviet Phys.—Solid State 4, 729 (1962)].} except near $\theta=0$, where (5) was used. The polarization of modes with **K** along the [110] direction is not known, except at one point of high symmetry in each zone of (periodically extended) reciprocal space, where the vibrations are purely longitudinal or transverse. Therefore, the measurement and calculations for the [110] direction are confined to angles near $\theta=0$.

was normal to the incident beam, **K** varied in direction at half the rate at which θ changed, when the analyzing slits were rotated. Figure 1 shows the measurements, corrected for background (angularly dependent), absorption, polarization, and slit-height smearing. The correction[15] for this last effect is a factor $[1+(b/a)\times(\theta^2-4.76)]/[1+(b/a)\theta^2]$, b/a in deg^{-2}, where the expression $a+b\theta^2$ is fitted to the experimental total-scattering curve at θ. The factor is 0.90 to 0.91 at $\theta=0$.

The experimental values for $I(0)$ are from 16 to 30% too large in the three directions. Only measurements made during the first month after the final etch are reported because the scattering is found to increase with time, most conspicuously at the smallest θ. The increase must represent deterioration, either surface corrosion or possibly x-ray damage. Scattering from surface roughness is known[17] to be prominent at small θ, especially in otherwise perfect crystals, where the volume scattering is small. Comparison of surface conditions shows that it accounts at least primarily for the deviation, at $\theta<4°$, of the experimental curves in Fig. 1 above the expected $a+b\theta^2$ variation. The observed excess in $I(0)$ is attributed to this cause. The time variation in $I(\theta)$ is approximately independent of θ, so that it is believed the final curve should be lowered uniformly by about 0.05 e.u./atom.

Similar experiments on quartz, in two directions, also show surface scattering, similar agreement with (5) (piezoelectric effects are negligible), and I_{inc} less than theory.[6] However, the phonon dispersion relations are not known.

[17] D. L. Weinberg, J. Appl. Phys. 33, 1012 (1962).

ACKNOWLEDGMENTS

The author thanks C. B. Walker for a helpful discussion, H. E. Hagy for helping to obtain the (110) Si, R. B. Emmons of Philco Corp. for providing it, and R. E. Szupillo for etching it.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Optically Active Phonon Processes in CdS and ZnS*

R. Marshall and S. S. Mitra
Physics Division, IIT Research Institute,† Chicago, Illinois
(Received 26 December 1963)

Transmission spectra of thin (10 to 30 μ) CdS single-crystal samples have been recorded in the hitherto uninvestigated region of 250 to 400 cm^{-1}. A number of new absorption maxima are located which necessitated a new phonon assignment with considerably lower values of the acoustic modes. Some 30 absorption peaks are attributed to multiphonon combinations of the following Brillouin zone-edge frequencies: LO = 295 cm^{-1}, TO$_1$ = 261 cm^{-1}, TO$_2$ = 238 cm^{-1}, LA = 149 cm^{-1}, TA$_1$ = 79 cm^{-1}, and TA$_2$ = 70 cm^{-1}. The transmission spectrum of hexagonal ZnS is obtained for the first time in the region 300 to 750 cm^{-1}. Again some 27 absorption maxima are assigned to multiphonon combinations of LO = 346 cm^{-1}, TO$_1$ = 318 cm^{-1}, TO$_2$ = 297 cm^{-1}, LA = 181 cm^{-1}, TA$_1$ = 92 cm^{-1}, and TA$_2$ = 73 cm^{-1}. The phonon assignment for the optical lattice absorption in cubic ZnS due to Deutsch is shown to be inconsistent with the requirements of the Brout sum rule and other regularities expected for the zincblende-type semiconductors. Two new assignments (LO = 339 cm^{-1}, TO = 298 cm^{-1}, LA = 155 cm^{-1}, and TA = 93 cm^{-1}, and LO = 339 cm^{-1}, TO = 298 cm^{-1}, LA = 190 cm^{-1}, and TA = 115 cm^{-1}) only slightly different from each other are proposed. The characteristic phonon frequencies are discussed in terms of structure and effective ionic charge.

1. INTRODUCTION

INFRARED lattice absorption arising from multiphonon combinations has been observed in many group III–V compounds and group IV elements.[1,2] The combination bands are a continuous absorption. However, peaks in absorption occur because of singularities in the phonon frequency distribution. These singularities arise from critical points in the Brillouin zone where the frequency wave vector ($\mathbf{k}$) curves have zero slope. Such critical points are believed to occur at or near the zone boundary. Usually the critical point characterizes a phonon branch and the combination bands may be classified into summations or differences of the longitudinal optical (LO), transverse optical (TO), longitudinal acoustic (LA) and transverse acoustic (TA) phonons at the zone edge. The interaction of an infrared photon with the lattice vibrations are governed only by the conservation of energy and wave vector. Turner and Reese[1] have assigned some forty components in the lattice absorption of AlSb in this manner.

Only a limited amout of experimental evidence of these effects has been obtained for group II–VI compounds. Mitra[3] has recently explained the infrared absorption observed in ZnSe by Aven *et al.*[4] in terms of phonon combinations, and it appears that the conditions are similar to those governing the behavior of III-V compounds. In the case of CdS, Collins[5] using the reflection technique located the zone-center optical transverse mode at 241 cm^{-1}. Subsequently, Balkanski and Besson[6] reported seven absorption maxima between 400 and 600 cm^{-1} which were assigned to various binary combinations using four assumed characteristic frequencies. Spitzer[7] has recently shown that the same two-phonon bands can also be explained by another set of four-phonon frequencies; one of which is different from those used by Balkanski and Besson. Deutsch[8] has reported a few additional bands, resolved under polarized radiation, in the same region. He proposed two different sets of assignments which almost fully explained his own observations but failed to explain some of the bands reported by Balkanski and Besson. Mitra[9] observed further absorption bands in the region 600–1400 cm^{-1} and used a five-frequency assignment for a comprehensive explanation. They were composed of the four frequencies of Balkanski and Besson and the alternative frequency of Spitzer.

All transmission measurements on CdS to date have been confined to frequencies greater than 400 cm^{-1}, despite the observation of Collins that the zone-center transverse optical-mode phonon has a value of 241 cm^{-1}. Thus, it was felt that a transmission study in the region below 400 cm^{-1} could provide additional important information, particularly regarding the low-frequency characteristic phonons. These were considerably larger than in the case of ZnSe[3] and the III–V compounds.[1] Furthermore, if the proposed frequencies are designated LO, TO, LA, and TA phonons, there is a lack of consistency with various correlations described by Keyes,[10] Mitra,[3] and Mitra and Marshall,[11] including the Brout sum rule.[12] It was also anticipated that by using very thin crystals, structure in the broad absorption bands would be resolved and be of considerable assistance in obtaining a reliable assignment.

* Supported by U. S. Air Force Cambridge Research Laboratories, Office of Aerospace Research, under Contract No. AF 19(628)-2418.

† Formerly Armour Research Foundation of Illinois Institute of Technology.

[1] W. J. Turner and W. E. Reese, Phys. Rev. 127, 126 (1962), and Refs. 1 to 7 therein.
[2] J. R. Hardy and S. D. Smith, Phil. Mag. 6, 1163 (1961).
[3] S. S. Mitra, Phys. Rev. 132, 986 (1963).
[4] M. Aven, D. T. F. Marple, and B. Segall, J. Appl. Phys. 32, 2261 (1961).
[5] R. J. Collins, J. Appl. Phys. 30, 1135 (1959).
[6] M. Balkanski and J. M. Besson, J. Appl. Phys. 32, 2292 (1961).
[7] W. G. Spitzer, J. Appl. Phys. 34, 792 (1963).
[8] T. Deutsch, J. Appl. Phys. 33, 751 (1962).
[9] S. S. Mitra, Phys. Letters 6, 249 (1963).
[10] R. W. Keyes, J. Chem. Phys. 37, 72 (1962).
[11] S. S. Mitra and R. Marshall (to be published).
[12] R. Brout, Phys. Rev. 113, 43 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

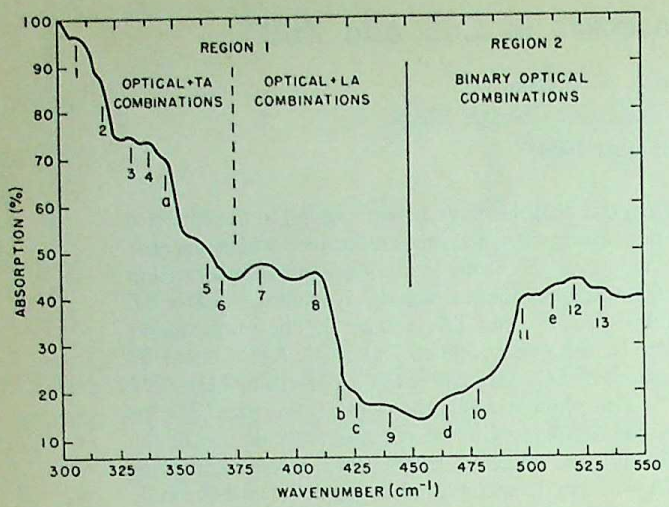


FIG. 1. Transmission spectrum of CdS in the region 300 to 550 cm^{-1}.

In the case of ZnS, Deutsch[13] has studied absorption in the cubic modification, over the frequency range 410–830 cm^{-1}. His assignment fails to explain the strong absorption at 431 cm^{-1}. Furthermore, the characteristic frequencies used (particularly the low-frequency value), as in the case of CdS, do not conform with the correlations mentioned earlier. Thus a new assignment for cubic ZnS, with lower values for the acoustic branches, seemed necessary.

Lattice absorption in hexagonal ZnS is reported here for the first time. Thin crystals have been used in order to cover the optical-acoustical phonon combination region and to resolve the structure involved in the anticipated removal of degeneracy of the zone-boundary TO mode and possibly the TA mode as well.

2. EXPERIMENTAL

Platelet crystals of CdS and ZnS, grown from the vapor phase by nucleation on a quartz seed as described by Fochs and Lunn,[14] were used in this investigation. This growth technique yields thin crystals ($\sim 10\ \mu$) of high optical quality and low impurity content, but relatively small surface areas ($<0.5\ cm^2$).

Absorption data was obtained for both compounds over the region 300–800 cm^{-1} using a Perkin Elmer double-beam spectrometer (model 521) with grating optics.

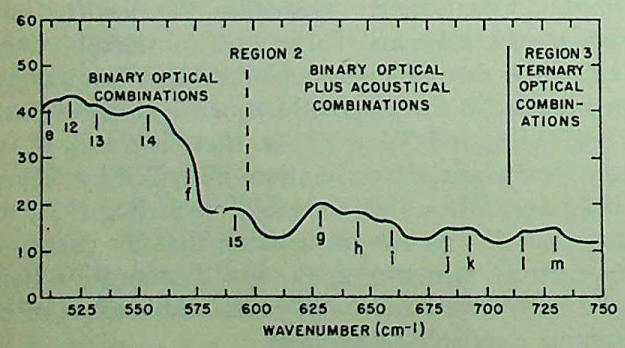


FIG. 2. Transmission spectrum of CdS in the region 525 to 750 cm^{-1}.

3. ANALYSIS OF SPECTRA

a. CdS

The absorption spectrum is shown in Figs. 1 and 2, and the positions of the bands observed listed in Table I. The absorption spectrum may be conveniently

TABLE I. Assignment of the absorption maxima observed in CdS in terms of phonon combinations of frequencies (cm^{-1}):

	$TO_1=261$		$TA_1=79$
$LO=295$		$LO=149$	
	$TO_2=238$		$TA_2=70$

Band No.	Present study (300°K) cm^{-1}	Balkanski and Besson (77°K) cm^{-1}	Deutsch (77°K) cm^{-1}	Assignment	Calcd. cm^{-1}	Calcd.–Obsd.
1	308			TO_2+TA_2	308	0
2	318			TO_2+TA_1	317	−1
3	331			TO_1+TA_2	331	0
4	340			TO_1+TA_1	340	0
a	347			TO_1+TO_2-LA	350	+3
5	365			$LO+TA_2$	365	0
6	374			$LO+TA_1$	374	0
7	387			TO_2+LA	387	0
		403				
8	410			TO_1+LA	410	0
b	421			$TO_1+TO_2-TA_1$	420	−1
c	428		432	$TO_1+TO_2-TA_2$	429	+1
9	442		443	$LO+LA$	444	+2
d	467			$TO_2+LA+TA_2$	466	−1
10	478			$2TO_2$	476	−2
11	498	500	502	TO_1+TO_2	499	−1
e	511			$2LO-TA_1$	511	0
12	522	524	526	$2TO_1$	522	0
13	533	540		$LO+TO_2$	533	0
14	555	562	562	$LO+TO_1$	556	+1
f	571	579	576	$TO_1+TO_2+TA_2$	569	−2
15	590	599	598	$2LO$	590	0
g	628			$LO+TO_1+TA_2$	626	−2
h	647			TO_1+TO_2+LA	648	+1
i	660			$2LO+TA_2$	660	0
j	684			$LO+TO_2+LA$	682	−2
k	695			$LO+TO_1+2TA_2$	696	+1
l	715			$3TO_2$	714	−1
m	730			$2LO+2TA_2$	730	0
	795a			$LO+TO_1+TO_2$	794	−1
	1045 (broad and weak)[a]			$LO+2TO_1+TO_2$	1055	+10
	1345 (broad and weak)[a]			$2LO+2TO_1+TO_2$	1350	+5

[a] Reported earlier (Ref. 9).

divided into three regions: (1) the group of strong bands below 450 cm^{-1}, (2) bands of medium absorption between 450 and 710 cm^{-1}, and (3) the weak bands above 700 cm^{-1}. Since the zone-center transverse optical-phonon frequency as determined by Collins is at 241 cm^{-1}, these three regions may be considered as arising from one, two, and three optical-phonon processes. First, considering the second region the six prominent bands [Nos. 10, 11, 12, 13 (Fig. 1) and Nos. 14 and 15 (Fig. 2)] are assigned to various binary combinations of the three zone-boundary optical phonons, $LO=295\ cm^{-1}$, $TO_1=261\ cm^{-1}$, and $TO_2=238\ cm^{-1}$. It may be commented here that the first two char-

[13] T. Deutsch, in *Proceedings of the International Conference on Semiconductors, Exeter, 1962* (The Institute of Physics and Physical Society, London, 1962), p. 505.
[14] P. D. Fochs and B. Lunn, J. Appl. Phys. 34, 1762 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

acteristic phonons have been considered by all the previous investigators, Balkanski and Besson, Deutsch, Spitzer, and Mitra. Since the zone-center value of the transverse phonon is 241 cm^{-1}, it seems reasonable that the zone-boundary phonons should split into 261 and 238 cm^{-1} in the hexagonal wurtzite structure.

Balkanski and Besson have studied the temperature dependence of the prominent bands in the two-optical-phonon region. In Table II, we have made a comparison, based on our assignment, between the calculated and observed values of the absorption coefficient at 80 and 293°K. It may be shown[15] that the net power absorbed in a two-phonon summation band $\nu_n+\nu_k$ is proportional to

$$\alpha_c=1+F_n+F_k, \quad (1)$$

where

$$F_n=[\exp(hc\nu_n/kt)-1]^{-1}, \quad (2)$$

and ν_n = wave number of the nth phonon. Thus the terms f_c and f_0 listed in Table II are

$$f_c=\alpha_c(80°K)/\alpha_c(293°K), \quad (3)$$

and

$$f_0=\alpha_0(80°K)/\alpha_0(293°K), \quad (4)$$

where α_0 is the observed absorption coefficient at the peak absorption after correction for background. The excellent agreement shown in Table II between the calculated and the observed temperature dependencies, although perhaps a bit fortuitous, is a further confirmation of the essential correctness of the assignment of the optical modes.

Having obtained the optical-mode frequencies, the acoustical-mode frequencies at the zone boundary are obtained from the prominent bands in region (1). A perusal of region (1) reveals that one may further subdivide it into two groups: one above and one below 375 cm^{-1}. The three bands at 387, 410, and 442 cm^{-1} (numbers 7, 8, and 9, respectively, in Fig. 1) seem an obvious choice for the bands arising from combinations of the three optical modes with the longitudinal acoustical mode at the zone boundary. Each one of them yields a value close to 149 cm^{-1} for LA. Below 375 cm^{-1} there

TABLE II. Comparison of the observed and predicted temperature dependence of the optical phonon combinations in CdS.

Peak position cm^{-1}	Assignment	Observed[a] f_0	Calculated f_c
498	TO_1+TO_2	0.56	0.556
522	$2TO_1$	0.57	0.574
533	$LO+TO_2$	0.60	0.578
555	$LO+TO_1$	0.64	0.598
590	2LO	0.63	0.624

[a] Determined from Balkanski and Besson's data (Ref. 6).

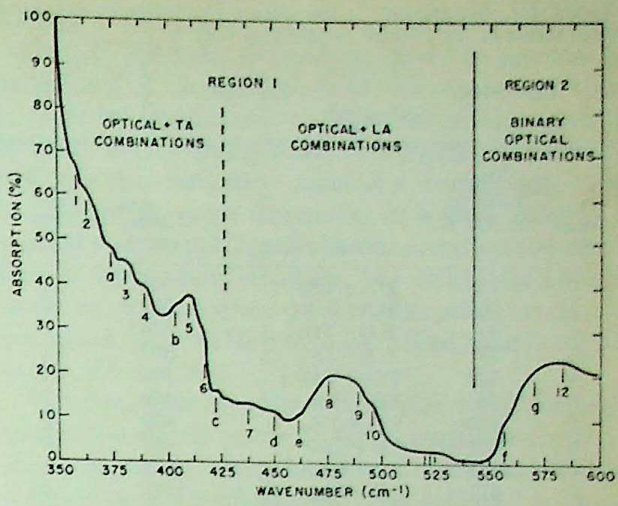


FIG. 3. Infrared absorption in wurtzite ZnS (350 cm^{-1} to 600 cm^{-1}).

are six prominent bands (Nos. 1 to 6 in Fig. 1). They occur as three pairs, each with an average splitting of 9 cm^{-1}. This splitting is attributed to two transverse acoustical phonons, which, like the transverse optical phonons, seem to be in order for the wurtzite structure. The two zone-boundary acoustical phonons assigned in this manner are $TA_1=79$ cm^{-1} and $TA_2=70$ cm^{-1}. The complete assignment for all the observed absorption maxima is presented in Table I.

b. Wurtzite ZnS

The absorption spectrum of wurtzite ZnS is shown in Figs. 3 and 4 and the band positions are listed in Table III. In this case, the absorption regions are not as clearly defined as in CdS, but similar broad classifications can be made: region 1—below 545 cm^{-1}, can further be divided into regions above and below 430 cm^{-1}; region 2—between 545 and 700 cm^{-1}, and region 3—above 700 cm^{-1}.

The absorption in region 2 is again thought to arise primarily from the binary combination of optical phonons with assumed characteristic phonon frequencies, $LO=346$ cm^{-1}, $TO_1=318$ cm^{-1}, and $TO_2=297$ cm^{-1}. In region 1 two areas are distinctive, comprising

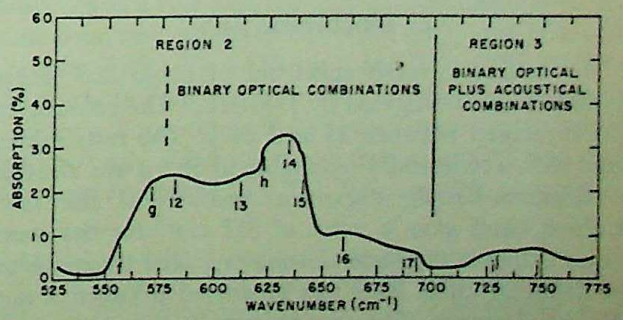


FIG. 4. Infrared absorption in wurtzite ZnS (525 cm^{-1} to 775 cm^{-1}).

[15] W. Cochran, S. J. Fray, J. A. Johnson, J. E. Quarrington, and N. Williams, J. Appl. Phys. 32, 2102 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Assignment of the lattice absorption in hexagonal ZnS in terms of the phonon frequencies (cm^{-1}):

	$TO_1=318$		$TA_1=92$
$LO=346$		$LA=181$	
	$TO_2=297$		$TA_2=73$

Band No.	Peak position at 300°K cm^{-1}	Assignment	Calcd. cm^{-1}	Calcd.–Obsd.
1	358	TO_2+TA_2	360	+2
2	362	2LA	362	0
a	374	$2LO-TO_1$	374	0
3	381	TO_2+TA_1	383	+2
4	390	TO_1+TA_2	391	+1
b	404	$TO_1+LA-TA_1$	406	−2
5	409[a]	TO_1+TA_1	409	0
6	419	$LO+TA_2$	419	0
c	423	$TO+LA-TA_2$	425	+2
7	437	$LO+TA_1$	437	0
d	450	$TO_1+TA_1+TA_2$	452	+2
e	463	TO_1+2TA_2	464	+1
8	475[a]	TO_2+LA	474	−1
9	487	$TO_1+TA_1+TA_2$	483	−4
10	498	TO_1+LA	500	+2
11	526	$LO+LA$	528	+2
f	555	$LO+TO_2-TA_1$	551	−4
g	571	$TO_1+LA+TA_2$	572	+1
12	583[a]	$2TO_2$	584	+1
13	612	TO_1+TO_2	610	−2
h	622	$LO+LA+TA_1$	620	−2
14	634[a]	$2TO_1$	636	+2
15	640[a]	$LO+TO_2$	638	−2
16	658	$LO+TO_1$	658	0
17	694	2LO	692	−2
i	730	$LO+TO_2+TA_1$	730	0
j	750	$2LO+TA_1$	750	0

[a] Relatively intense bands.

the absorption maxima arising from the combination of the optical phonons with the zone boundary TA (between 370 and 450 cm^{-1}) and with the zone boundary LA (between 435 and 525 cm^{-1}) phonons. Six components are observed in the optical-plus-TA region, and again it has been necessary to assume that the TA modes are nondegenerate. Characteristic phonon frequencies of $TA_1=92$ cm^{-1} and $TA_2=73$ cm^{-1} predict the observed structure in the absorption spectrum with good accuracy. From the optical-plus-LA combination bands (8, 10, 11, Fig. 3), a value of $LA=181$ cm^{-1} has been obtained. This also explains the relatively intense absorption band observed at 362 cm^{-1} as a 2LA combination.

c. Zincblende ZnS

The lattice infrared spectrum of cubic ZnS single crystals was investigated by Deutsch.[13] The reflectivity was measured between 15 and 34 μ. The reststrahlen band with a reflectivity maximum at 30.8 μ was located. A Kramers-Kronig dispersion analysis of the reststrahlen band gave a value of 312 cm^{-1} for the zone-center transverse optical phonon (ν_t). The transmission spectrum of cubic ZnS was measured between 1 and 24 μ by Deutsch. The transmission is essentially constant up to 14 μ. Ten absorption maxima were observed at longer wavelengths between 430 and 820 cm^{-1}. These bands were attributed to optically active two- and three-phonon processes. Using characteristic zone-boundary phonon frequencies of $A=379$ cm^{-1}, $B=297$ cm^{-1}, $C=263$ cm^{-1}, and $D=228$ cm^{-1}, Deutsch could explain seven out of the ten observed bands.

Since a diatomic cubic crystal usually has two optical and two acoustic branches, the higher two phonons must be from optical and lower two from acoustic branches. In a predominantly ionic crystal like ZnS, one would expect that the LO modes should have a higher frequency than the TO modes throughout the Brillouin zone. One may thus identify the four characteristic phonon frequencies, A, B, C, and D assigned by Deutsch as zone boundary LO, TO, LA, and TA, respectively. However, a comparison[3] with the phonon frequencies of the other zincblende-type crystals reveals that the acoustical phonon frequencies as assigned by Deutsch are unusually large. Furthermore, since the zone-center longitudinal optical phonon frequency (ν_l) is 377 cm^{-1} as calculated from the Lyddane-Sachs-Teller formula,[16] a value of 379 cm^{-1} for the zone boundary LO seems improbable. The Brout sum at the zone center amounts to 1.200×10^{28} sec^{-2} as compared with a value of 1.754×10^{28} sec^{-2} near the zone edge. This large difference (45%) is quite unexpected for a predominantly ionic cubic crystal such as ZnS, and may be attributed to high values of LA and TA phonons as assigned by Deutsch. The discrepancy is even larger if one uses the ν_t given by Mitsuishi, Yoshinaga, and Fuyita.[17]

A fresh assignment of the zone-boundary characteristic phonon frequencies of cubic ZnS with lower values of the acoustic modes has therefore been attempted.

TABLE IV. Assignment of the multiphonon absorption bands in cubic ZnS.

	Deutsch			Present study		
Observed[a] cm^{-1}	Assignment[f]	Calc. (cm^{-1})	Assignment 1[g]	Calc. (cm^{-1})	Assignment 2[h]	Calc. (cm^{-1})
431[b]	...	...	LO+TA	432	TO+TA	413
415[c]						
455	2D	456	TO+LA	453	LO+TA	454
491[b]	C+D	491	LO+LA	494	TO+LA	488
526	2C, B+D	526	LO+2TA	525	LO+LA	529
593[b]	2B	594	2TO	596	2TO	596
605[d]	...	...	TO+2LA	608	LO+LA+TA	603
642[b]	A+C	642	LO+TO	637	LO+TO	637
677[b]	A+B	677	2LO	678	2LO	678
733[c]	...	...	LO+TO+TA	731	LO+2LA	719
823	2B+D	822	LO+TO+2TA	824	LO+TO+LA	827

[a] Data from Deutsch (Ref. 13).
[b] Relatively stronger bands.
[c] Although Deutsch lists a value of 431 cm^{-1}, this band appears to be beyond the range of his observation. A lower value seems probable both from the extrapolation of his spectrum and from a comparison of his spectrum with that of the hexagonal modification reported in this paper.
[d] Not listed by Deutsch but is apparent from his spectra.
[e] Broad and very weak.
[f] $A=379$ cm^{-1}, $B=297$ cm^{-1}, $C=263$ cm^{-1}, and $D=228$ cm^{-1}.
[g] LO=339 cm^{-1}, TO=298 cm^{-1}, LA=155 cm^{-1}, and TA=93 cm^{-1}.
[h] LO=339 cm^{-1}, TO=298 cm^{-1}, LA=190 cm^{-1}, and TA=115 cm^{-1}.

[16] R. H. Lyddane, R. G. Sachs, and E. Teller, Phys. Rev. 59, 673 (1941).
[17] A. Mitshuishi, H. Yoshinaga, and S. Fuyita, J. Phys. Soc. Japan 13, 1235 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE V. Comparison of the observed and predicted temperature dependence of the two-phonon combinations in cubic ZnS.

Peak position cm^{-1}	Assignment	Observed[a] $\alpha(77°K)/\alpha(300°K)$	Observed[a] $\alpha(420°K)/\alpha(300°K)$	Calculated $\alpha(77°K)/\alpha(300°K)$	Calculated $\alpha(420°K)/\alpha(300°K)$
491	LO+LA[b]	0.50	...	0.490	1.34
	TO+LA[c]	0.50	...	0.518	1.34
593	2TO	0.69	1.25	0.616	1.31
642	LO+TO	0.79	1.14	0.643	1.29
677	2LO	0.80	1.12	0.671	1.28

[a] Determined from Deutsch's spectra (Ref. 13).
[b] Assignment 1.
[c] Assignment 2.

With assumed values of LO = 339 cm^{-1}, TO = 298 cm^{-1}, LA = 155 cm^{-1}, and TA = 93 cm^{-1}, all the observed absorption maxima could be satisfactorily explained. However, the LA-phonon frequency seems to be a bit low when compared with the corresponding phonon frequency of the hexagonal ZnS. An alternative assignment with the same LO and TO but LA = 190 cm^{-1} and TA = 115 cm^{-1} seems also possible although the numerical agreement with the observed band positions is a trifle worsened in this case. The assignments are shown in Table IV.

Deutsch's data are used for the comparison of observed and calculated temperature dependences. Table V shows the calculated and observed temperature factors for the relatively prominent bands. There are usually two contributions to the background absorption; the first due to free charge carriers and the second to the tail of the resonance absorption at ν_t. The former is expected to decrease with the energy of the absorbed photon. Since the background absorption in the high-frequency side of the lattice bands of ZnS is very low and independent of temperature, the contribution of free carrier absorption to the background may be completely neglected. This is to be expected for a crystal of large band gap. The essentially temperature-independent contribution from ν_t to the background absorption is obtained from the extinction coefficient calculated from the reststrahlen spectrum (the temperature-dependent portion chiefly arises from the multiphonon processes). For the bands where data is available, the absorption coefficient is rather temperature insensitive. Thus the agreement between the observed and the calculated values may be regarded as only fair. The two assignments proposed here only differ in the case of the band at 491 cm^{-1}. Temperature dependence alone can not therefore decide between the two assignments. A fresh high-resolution transmission study of the cubic sample seems desirable.

4. DISCUSSION

As noted earlier, the multiple structure observed in the lattice infrared spectra of the zincblende-type crystals can usually be accounted for in terms of four characteristic phonon energies belonging to the four branches: TA. LA, TO, and LO at the Brillouin zone boundary. In certain cases, notably diamond[2] and GaAs,[15] it has been deemed necessary to assume the breakdown of the degeneracy of the transverse modes in order to satisfactorily account for all the observed optically active multiphonon combinations. Since the Brillouin zone-boundary value of a branch may differ in different $\mathbf{k}$ space directions, it is only natural to assume such a splitting of the transverse modes in the case of the wurtzite structure. For satisfactory assignments of the numerous lattice absorption maxima in hexagonal CdS and ZnS we have indeed taken recourse to two TA and two TO phonons.

The transverse frequency is smaller than the longitudinal one for all values of the propagation vector for the acoustical branches, except at the zone center where both are zero. Therefore there is no difficulty in identifying the lowest characteristic phonon frequencies with definite TA and LA modes. However, the identification of the optical-branch phonons to definite transverse and longitudinal modes is not so clear cut, and some confusion has resulted in the past.

Keyes[19] has noted some correlation between the zone-boundary optical-phonon energies of the zincblende-type crystals and the effective ionic charge q^*, derived from the Szigeti relation.[18] For a purely covalent crystal ($q^*=0$) such as diamond or Ge, the transverse and longitudinal optical branches are degenerate at the zone center whereas TO > LO at the zone boundary. Although $\nu_1 > \nu_t$ at $\mathbf{k}=0$ (Lyddane-Sachs-Teller formula[16]) for weakly ionic crystals like InSb ($q^*=0.42$), TO is still greater than LO at the zone boundary. In the case of a crystal of intermediate ionicity ($q^* \simeq 0.7$), Keyes predicted that LO = TO at $\mathbf{k}_{max}$, which indeed is the case with ZnSe as shown by Mitra.[3] However, strongly ionic crystals ($q^* > 0.7$) such as the alkali halides[19] are expected to have LO greater than TO for all values of $\mathbf{k}$.

The effective ionic charge on CdS and ZnS is 1.00 and 0.854, respectively. The highest characteristic phonon frequency for these materials may therefore be termed the zone boundary LO. The next two characteristic frequencies of the hexagonal CdS and ZnS, and the next single characteristic frequency of the cubic ZnS may now be identified with the TO modes. For predominantly ionic crystals, the transverse optical branches as a rule undergo only limited dispersion. The zone boundary values of $TO_1 = 261$ cm^{-1} and $TO_2 = 238$ cm^{-1} for CdS, $TO_1 = 318$ cm^{-1} and $TO_2 = 297$ cm^{-1} for the wurtzite ZnS and TO = 298 cm^{-1} for the zincblende ZnS are not too different from their zone-center values, 241 cm^{-1} and 312 cm^{-1}, respectively, for CdS and ZnS. The zone boundary LO, TO, LA, and TA phonon fre-

[18] B. Szigeti, Trans. Faraday Soc. 45, 155 (1949).
[19] A. D. B. Woods, W. Cochran, and B. N. Brockhouse, Phys. Rev. 119, 980 (1960); A. D. B. Woods, B. N. Brockhouse, R. A. Cowley, and W. Cochran, *ibid.* 131, 1025 (1963); R. A. Cowley, W. Cochran, B. N. Brockhouse, and A. D. B. Woods, *ibid.* 131, 1030 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

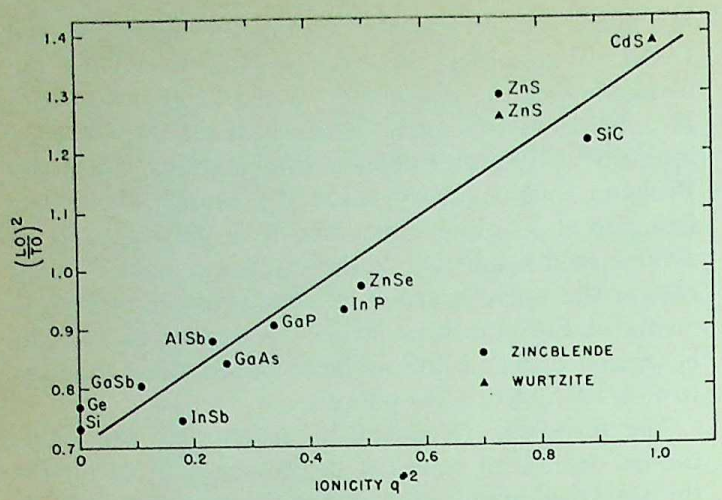


FIG. 5. Correlation of the ratio of the optical frequencies of the zincblende- and wurtzite-type semiconductors with the effective ionic charge.

quencies for a number of IV-IV, III-V and II-VI compounds along with their effective ionic charge are listed in Table VI. Keyes noted that $(LO/TO)^2$ is a linear function of q^{*2} for crystals of diamond and zincblende-type structures. Such a plot is shown in Fig. 5. The crystals of wurtzite structure also seem to follow the general pattern if one uses the mean of the two TO values.

The Brout sum rule[20] for lattice vibrations,

$$\sum_i \omega_i^2(\mathbf{k}) = \text{constant}, \qquad (5)$$

where $\omega_i(\mathbf{k})$ is the circular frequency in the ith branch when the wave vector is $\mathbf{k}$, was believed to be strictly valid for diatomic ionic crystals of cubic symmetry,

TABLE VI. Zone-boundary phonon frequencies and effective ionic charge of some IV-IV, III-V, and II-VI semiconductors.[a]

Compound	LO cm^{-1}	TO cm^{-1}	LA cm^{-1}	TA cm^{-1}	q^*
C[b]	1161	1275	991	750	0
Si	413	482	333	139	0
Ge[c]	247	280	215	65	0
SiC	851	770	540	363	0.94
GaSb	193	215	134	49	0.33
InSb	155	179	118	43	0.42
GaAs	234	258	188	70	0.51
AlSb	297	316	132	65	0.53
GaP	361	378	197	115	0.58
InP	318	329	150	62	0.68
ZnSe	212	208	162	87	0.71
ZnS(cubic)[d]	339	298	155	93	0.85
	339	298	190	115	
ZnS(hexagonal)[e]	346	318; 297	181	92; 73	
CdS(hexagonal)[f]	295	261; 238	149	79; 70	1.00

[a] Data from Ref. 3 if not otherwise stated.
[b] Reference 2.
[c] B. N. Brockhouse and P. K. Iyengar, Phys. Rev. 111, 747 (1958).
[d] This study (assignment 1).
[e] This study (assignment 2).
[f] This study.

[20] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press Inc., New York, 1963), p. 115.

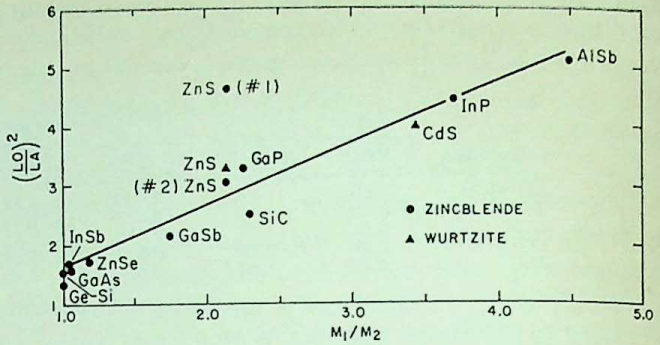


FIG. 6. The effect of mass difference on the separation between the optical and the acoustical branches.

whose constituent atoms interact only by Coulomb forces between all ions and the repulsive interaction is between nearest neighbors alone. However, relation (5) seems to hold reasonably well for a number of zincblende[1,3]-type crystals. Rosenstock[21] has recently demonstrated that Eq. (5) with little or no modification can also be applied to more general force models and to nonionic crystals.

In Table VII the zone-edge Brout sums $(\nu_1^2+2\nu_t^2)$ for CdS and ZnS are compared with their zone-boundary values $(LO^2+2TO^2+LA^2+2TA^2)$. The agreement is fair. Such a comparison may therefore be regarded as a check on the upper bounds of the zone-boundary phonon frequencies of diatomic crystals.

Other than the numerical agreement between the observed and calculated positions of infrared absorption peaks involving the acoustical modes, there are but few checks on the soundness of the LA and TA assignments. In analogy with the diatomic linear chain, Keyes has shown that one may expect a linear relationship between $(LO/LA)^2$ and σ, the ratio of the larger to the smaller ionic masses for the zincblende-type crystals. Such a plot is shown in Fig. 6, in which the wurtzite crystals are included as well. The LA assignment of hexagonal CdS and ZnS seems reasonable. For the cubic ZnS assignment, No. 2 seems to be more probable. If such is the case, the difference between the hexagonal and cubic ZnS characteristic phonon frequencies is not very

TABLE VII. Comparison of the zone-center and zone-edge Brout sums for some 2–6 compounds.

Compound	Structure	Brout sum at $\mathbf{k}\approx 0$ in 10^{27} sec^{-2}	Brout sum at $\mathbf{k}_{max}$ in 10^{27} sec^{-2}
CdS	Wurtzite	7.73	8.70
ZnS	Zincblende	12.00	11.84[a]
			12.60[b]
ZnS	Wurtzite	...	12.61
ZnSe	Zincblende	5.35	6.13

[a] Assignment 1.
[b] Assignment 2.

[21] H. B. Rosenstock, Phys. Rev. 129, 1959 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

large, which is to be expected from the essential similarity in the gross features of the spectra of these two modifications. Such a similarity was also noted between cubic and hexagonal modifications of SiC by Spitzer, Kleinman, and Walsh.[22]

[22] W. G. Spitzer, D. Kleinman, and D. Walsh, Phys. Rev. 113, 127 (1959); W. G. Spitzer, D. A. Kleinman, and C. J. Frosch, *ibid.* 113, 133 (1959).

ACKNOWLEDGMENTS

The authors are grateful to Dr. S. Nudelman and Dr. P. J. Gielisse of the Cambridge Research Laboratories for their interest in the investigation. Thanks are also due to Dr. P. D. Fochs, A. E. I. Research Laboratories, Harlow, England and Dr. B. Kulp of Aeronautical Research Laboratories for providing the specimens.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Theory of Electron Backscattering*

ROGER F. DASHEN†
Sandia Laboratory, Albuquerque, New Mexico
(Received 23 December 1963)

An exact integral equation is derived for the backscattering coefficient as a function of energy and angle. The equation has an approximate symmetry which allows one to determine the conditions under which backscattering will be independent of the energy of the incident electrons. The validity of some approximations is discussed, and a possible method for obtaining numerical solutions is proposed. It is found that small and large-angle individual scatterings are of about equal importance. For those electrons that have lost very little energy, the equation can be solved exactly, and some interesting results are obtained.

1. INTRODUCTION

WHEN a beam of electrons is directed against a thick solid target, many of the electrons stop in the target but some are scattered back out. These are the backscattered electrons.[1] Over the past few years a considerable amount of experimental information has been collected [2–6] and from an empirical point of view there now exists a fairly complete picture of the phenomenon. On the other hand, attempts at a theoretical interpretation of the data[7–10] have been based on rather severe approximations and have met with only limited success. The purpose of this paper is to present a reformulation of the problem which circumvents many, but by no means all of the difficulties which one encounters in trying to treat backscattering through the usual methods of multiple-scattering theory.

Mathematically, the problem can be formulated as follows. Given a uniform, monoenergetic flux of electrons incident on a semi-infinite block of scattering material, what is the energy and angular distribution of electrons that are scattered back out? In order to solve this problem using standard multiple-scattering theory, one must find the distribution function for the electrons at each point within the block. A more economical method would be to obtain an equation for the distribution of backscattered electrons alone, thereby freeing ourselves from the unnecessary task of finding the distribution within the block. In the following section we will derive an integral equation for the backscattering coefficient as a function of energy and angle. An application of this equation to the energy dependence of the backscattering coefficient is given in Sec. 3. Section 4 contains a critical analysis of some

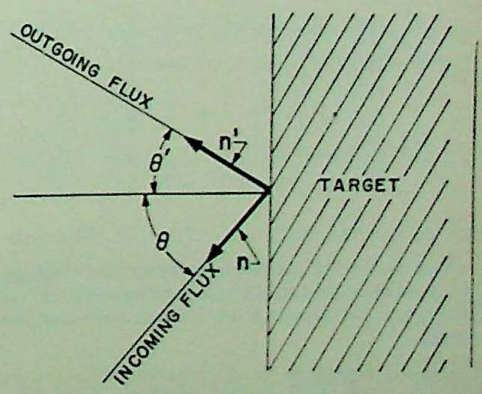


FIG. 1. Kinematics.

* Work was performed under the auspices of the U. S. Atomic Energy Commission.
† Present address: Department of Physics, California Institute of Technology, Pasadena, California.
[1] Secondary electrons are also emitted. Experimentally, it is usually assumed that an electron emerging from the target with an energy of more than 50 eV has been backscattered.
[2] E. Stenglass, Phys. Rev. 95, 345 (1954).
[3] P. Palluel, Compt. Rend. 224, 1492 (1947).
[4] H. Kanter, Ann. Physik 20, 144 (1957).
[5] K. Wright and J. Trump, J. Appl. Phys. 33, 687 (1962).
[6] H. Kulenkampff and K. Rüttiger, Z. Physik 137, 426 (1954).
[7] H. Bethe, Ann. Physik 5, 325 (1940).
[8] G. Archard, J. Appl. Phys. 32, 1505 (1961).
[9] T. Everhart, J. Appl. Phys. 31, 1483 (1960).
[10] W. Bothe, Z. Naturforsch. 4a, 542 (1949).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

approximate methods along with some interesting numerical results based on the assumption that small-angle scattering is unimportant. Finally, in Sec. 5 we discuss the possibility of more extensive calculations.

2. THE BASIC EQUATION

Given a uniform flux of electrons incident on a plane surface (Fig. 1), we define the backscattering coefficient R to be

$$R(\mathbf{n},E;\mathbf{n}',E')=\frac{\text{outgoing flux per unit solid angle per unit energy interval}}{\text{incident flux}}, \tag{1}$$

where E=incident energy, E'=outgoing energy, $\mathbf{n}$=unit vector *antiparallel* to the incoming direction, $\mathbf{n}'$=unit vector *parallel* to the outgoing direction. One will note that the quantity measured in the laboratory is not R, but $R(\mathbf{n},E,\mathbf{n}'E')\sec\theta\cos\theta'$, where θ is the angle between the $\mathbf{n}$ and the normal to the surface. This is due to the fact that we have assumed a uniform, infinitely broad, incoming beam whereas in the laboratory the incoming beam is confined to a thin pencil.

To derive an integral equation for R, we will follow a chain of thought that has proven to be very fruitful in similar contexts and generally goes under the title of "invariant imbedding." Our application of these ideas will be of the most rudimentary nature. For a thorough exposition of this manner of thinking and its many applications, the interested reader is referred to the review article of Bellman, Kalaba, and Wing.[11] The basic idea is very simple. Suppose we add a thin layer of the same material to the surface of our target thereby increasing its thickness by an amount Δt. Since we have assumed that our target was already infinitely thick, the backscattering coefficient must remain unchanged. If we take the new layer to be differentially thin, this simple device will lead to an integral equation for R.

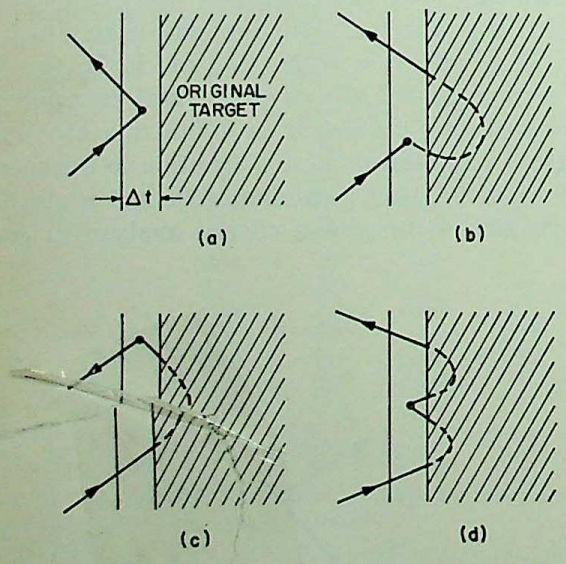


FIG. 2. New paths created by increasing the thickness of the target. The dots represent a single elastic or inelastic collision in the new layer, and the broken lines indicate backscattering in the original target.

[11] R. Bellman, R. Kalaba, and G. Wing, J. Math. Phys. 1, 280 (1960).

As notational difficulties would tend to obscure the basic simplicity of the argument, we will not follow the derivation in mathematical detail. The reader, if so inclined, can easily fill in the gaps.

Apart from normalization factors, the backscattering coefficient $R(\mathbf{n},E;\mathbf{n}',E')$ is imply the probability that an electron will follow any path which can lead to backscattering from the state $(\mathbf{n},E)$ to the state $(\mathbf{n}',E')$. By adding a layer of scattering material we have created some new paths. These are the paths where the electron makes a collision, either elastic or inelastic, somewhere in the new layer. They are illustrated schematically in Fig. 2. Note that we need not consider paths where the electron makes two or more collisions in the added layer since the probability that an electron will follow one of these paths is of order $(\Delta t)^2$. To calculate the *increase* in R due to fact that we have added some extra paths, we must add up the probabilities that an electron will follow each of the new paths. Referring to Fig. 2 we see that the sum over paths of type (a) is simply a normalization factor times the atomic cross section. Again, apart from normalization factors, the paths of type (b) can be summed by multiplying the atomic cross section for scattering from the state $(E,\mathbf{n})$ to a state $(E'',\mathbf{n}'')$ times the "original" backscattering coefficient $R(\mathbf{n}'',E'';\mathbf{n}',E')$ and summing over all intermediate states $(\mathbf{n}'',E'')$. Paths of type (c) are treated similarly. The sum over paths of type (d) contains the atomic cross section times the square of the "original" backscattering coefficient and a double summation over intermediate states.

Since the total variation in R must be zero, the increase in probability of backscattering due to the increased number of available paths must be exactly cancelled by a decrease in the probability that an electron will follow one of the old paths; i.e., a path along which it makes no collisions in the added layer. It is immediately obvious that the net probability that an electron will follow one of the original paths is decreased since the probability that an electron will pass twice through the new layer without making a collision is less than one. More specifically, the *reduction* in the probability that an electron will follow one of the old paths is simply the "original" backscattering coefficient times the total probability that the electron will make a collision as it passes twice through the new layer.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

At this point we make the usual assumption that inelastic collisions with the atomic electrons contribute only to energy loss and do not change the incident electrons' direction of motion. This simplification is not in principle necessary, but it is a very good approximation and has the pleasant feature that the cross sections for angular deflection and energy loss become independent. Combining this assumption with the analysis of the preceding paragraphs, one finds that R satisfies the integral equation

$$\begin{aligned}&[(\sigma_T(E)+\omega_T(E))\sec\theta+(\sigma_T(E')+\omega_T(E'))\sec\theta']\\&\qquad\times R(\mathbf{n},E;\mathbf{n}',E')\\&=\sigma(E,-\mathbf{n}\cdot\mathbf{n}')\sec\theta'\delta(E-E')\\&\quad+\int_{E'}^{E}[\omega(E,E'')\sec\theta R(\mathbf{n},E'';\mathbf{n}',E')\\&\quad+\omega(E'',E')\sec\theta' R(\mathbf{n},E;\mathbf{n}',E'')]dE''\\&\quad+\int[\sigma(E,\mathbf{n}\cdot\mathbf{n}'')\sec\theta'' R(\mathbf{n}'',E;\mathbf{n}',E')\\&\quad+\sigma(E',\mathbf{n}''\cdot\mathbf{n}')\sec\theta' R(\mathbf{n},E;\mathbf{n}'',E')]d\mathbf{n}''\\&\quad+\int_{E'}^{E}\int\int R(\mathbf{n},E;\mathbf{n}'',E'')\sigma(E'',-\mathbf{n}''\cdot\mathbf{n}''')\\&\qquad\times\sec\theta''' R(\mathbf{n}''',E'';\mathbf{n}',E')dE''d\mathbf{n}''d\mathbf{n}''',\end{aligned}\tag{2}$$

where $\sigma(E,\cos\Theta)$ is the atomic cross section for elastic scattering through an angle Θ, $\omega(E,E')$ is the atomic cross section for an inelastic collision leading to an energy loss $E-E'$, $\sigma_T(E)$ and $\omega_T(E)$ are the total elastic and inelastic cross sections, $\sec\theta$ is the secant of the angle between $\mathbf{n}$ and the normal, and the angular integrations run over the hemisphere of unit vectors pointing out of the target. The first term on the right of Eq. (2) is the probability that an electron will follow one of the new paths of type (a). In accord with our assumption that inelastic collisions do not cause angular deflection, we have included only the elastic cross section in this term. The two single integrals are the sum over paths of types (b) and (c) and the triple integral is the sum over paths of type (d). Finally, the left-hand side is the net reduction in the probability that an electron will follow one of the old paths. The geometrical factors $\sec\theta$ are the usual ones which arise because the effective volume of the layer available for scattering into a direction $\mathbf{n}$ is proportional to $\Delta t\sec\theta$. In writing down Eq. (2) we have omitted a factor $\Delta t\times$(number of atoms per unit volume) which multiplies each term as it is of no consequence. Physically, this means that backscattering is independent of the density of the scattering material.

For future purposes it is convenient to write Eq. (2) in the form

$$\begin{aligned}&\int_{E'}^{E}\int[\delta(E-E'')\omega_T(E)-\omega(E,E'')]\\&\qquad\times\sec\theta''\delta(\mathbf{n}-\mathbf{n}'')R(\mathbf{n}'',E'';\mathbf{n}',E')dE''d\mathbf{n}''\\&\quad+\int_{E'}^{E}\int R(\mathbf{n},E;\mathbf{n}'',E'')[\delta(E''-E')\omega_T(E')\\&\qquad-\omega(E'',E')]\sec\theta'\delta(\mathbf{n}'-\mathbf{n}'')dE''d\mathbf{n}''\\&=\sigma(E,-\mathbf{n}\cdot\mathbf{n}')\sec\theta'\delta(E-E')\\&\quad+\int_{E'}^{E}\int[[\sigma(E,\mathbf{n}\cdot\mathbf{n}'')-\delta(\mathbf{n}-\mathbf{n}')\sigma_T(E)]\\&\qquad\times\sec\theta''\delta(E-E'')R(\mathbf{n}'',E'';\mathbf{n}',E')]dE''d\mathbf{n}''\\&\quad+\int_{E'}^{E}\int[R(\mathbf{n},E;\mathbf{n}'',E'')[\sigma(E'',\mathbf{n}'\cdot\mathbf{n}'')\\&\qquad-\delta(\mathbf{n}'-\mathbf{n}'')\sigma_T(E'')]\sec\theta'\delta(E''-E')]dE''d\mathbf{n}''\\&\quad+\int_{E'}^{E}\int_{E'}^{E''}\int\int R(\mathbf{n},E;\mathbf{n}'',E'')\\&\qquad\times\sigma(E'',-\mathbf{n}''\cdot\mathbf{n}''')\sec\theta'''\delta(E''-E''')\\&\qquad\times R(\mathbf{n}''',E''';\mathbf{n}',E')dE''dE'''d\mathbf{n}''d\mathbf{n}''',\end{aligned}\tag{3}$$

where $\delta(\mathbf{n}-\mathbf{n}')$ is the delta function on the unit sphere.

In the remaining sections of this paper we will restrict ourselves to electron energies that are nonrelativistic but large compared to the binding energy of the atomic electrons. This is a permissible limitation since from the form of Eq. (2) it is apparent that for a given maximum value of E and minimum value of E', R does not depend on reflection coefficients whose energy indices lie outside this range. Under these conditions the elastic cross section is simply the Rutherford cross section[12]

$$\sigma(E,\cos\Theta)=\frac{Z^2e^4}{4E^2}\frac{1}{(1-\cos\Theta+\gamma)^2},\tag{4}$$

where γ is the screening parameter. The inelastic cross section can be sufficiently well approximated by[13]

$$\omega(E,E')=\frac{\pi e^4 Z}{E}\frac{1}{(E-E')^2},\quad \epsilon_{\max}>E-E'>\epsilon_{\min}\tag{5}$$

where $\epsilon_{\min}$ and $\epsilon_{\max}$ are the minimum and maximum energy losses.[13] For our purposes it is sufficient to know that $\epsilon_{\max}$ is proportional to E and that the usual range-

[12] N. Mott and H. Massey, *Theory of Atomic Collisions* (Oxford University Press, New York, 1949).

[13] M. Livingston and H. Bethe, Rev. Mod. Phys. 2, 245 (1937).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

energy relation is given by[12]

$$\int_0^E \omega(E,E')(E-E')dE' = -\frac{dE}{dx} = \frac{2\pi Ze^4}{E}\ln\left(\frac{2E}{I}\right), \quad (6)$$

where the unit of length has been chosen so that there is one atom per unit volume and I is a mean ionization potential.

In writing down Eq. (3) we have placed the factors in a particular order to bring out the fact that the integrations over energy and angle can be considered as matrix multiplications. This observation leads to a more economical notation and allows us to use the methods of matrix analysis to manipulate the equation. We consider $R(\mathbf{n},E;\mathbf{n}',E')$ as a matrix with two sets of indices, $\mathbf{n}$ and E, and define the following matrices

$$S_B = 4(e^4Z^2)^{-1}\sigma(E, -\mathbf{n}\cdot\mathbf{n}')\sec\theta'\delta(E-E'), \quad (7a)$$

$$S_F = 4(e^4Z^2)^{-1}[\sigma(E,\mathbf{n}\cdot\mathbf{n}')-\delta(\mathbf{n}-\mathbf{n}')\sigma_T(E)] \times\sec\theta'\delta(E-E'), \quad (7b)$$

$$W = (2\pi e^4Z)^{-1}[\delta(E-E')\omega_T(E)-\omega(E,E')] \times\sec\theta'\delta(\mathbf{n}-\mathbf{n}'). \quad (7c)$$

Equation (2) can now be written as

$$WR+RW = (Z/8\pi)(S_B+S_FR+RS_F+RS_BR). \quad (8)$$

Owing to the fact that none of the matrices S_B, S_F, and W commute, Eq. (8) cannot be solved explicitly. We can, however, put it in a more advantageous form. It is a standard theorem of matrix calculus[14] that a formal solution to (8) is given by

$$R = \frac{Z}{8\pi}\int_0^\infty \exp(-Wt)(S_B+S_FR+RS_F+RS_BR) \times\exp(-Wt)dt, \quad (9)$$

which holds provided that the integral converges. To see that the integral does converge one need only consider the physical meaning of the exponentials. Writing the matrix $\exp(-tW)$ in the form

$$\exp(-tW) = P(E,E';t\sec\theta)\delta(\mathbf{n}-\mathbf{n}'), \quad (10)$$

we find that P satisfies the differential equation and boundary condition

$$\frac{d}{d(t\sec\theta)}P(E,E';t\sec\theta) = (2\pi e^4Z)^{-1}\int_{E'}^{E}[\omega(E,E'')-\delta(E-E'')\omega_T(E)] \times P(E'',E',t\sec\theta)dE'', \quad (11)$$

$$P(E,E';0) = \delta(E-E'),$$

from which one easily sees that $P(E,E';t\sec\theta)$ is simply the probability that an electron will suffer an energy loss $E-E'$ while traversing a path of length $t\sec\theta$ where the unit of distance is chosen such that $2\pi e^4Z \times$(number of atoms per unit volume)$=1$. Now for fixed E and E', P tends rapidly to zero for large t and the integral converges. Landau[15] has given explicit expressions for the function P so we now have an equation that can, at least in principle, be solved by successive approximations.

Equations (8) and (9), being quadratic in R, most likely have more than one solution. Obviously the physically interesting solution has the property that, considered as a function of Z, it tends smoothly to zero for small Z. We will now give a heuristic proof that Eq. (9) has a unique solution with this property. Strictly speaking, the atomic screening radius γ and the minimum energy loss $\epsilon_{\min}$ are functions of Z. In the next section, however, we will see that the dependence of R on these quantities is rather unimportant and to a good approximation we can take the matrices S_F, S_B, and W to be independent of Z. With this assumption one can easily see that Eq. (9) has a unique formal solution as a power series in Z. To see that this series is, in fact, convergent we need only observe that an expansion of R in powers of Z corresponds exactly to an expansion in the number of times an individual electron is elastically scattered before leaving the target. Now a large number of scatterings requires a long time spent in the target, but for a given energy loss very long paths are highly improbable. Therefore, for fixed $E-E'$, the coefficient of Z^N must tend to zero for large N and the series converges. Since we have found that (8) has a unique solution which is analytic around $Z=0$, any extraneous solutions must have some singular behavior for small Z.

3. DEPENDENCE ON THE INCIDENT ENERGY

Perhaps the most striking experimental observation is that the number of backscattered electrons and their distribution in energy and angle are, over a broad range of energies, very nearly independent of the incident energy. The first application of our formalism will therefore be to investigate the energy dependence of R. Since the parameters γ and $\epsilon_{\min}$ appearing in (5) are rather poorly known functions of E and Z, our first task must be to see if R is sensitive to moderate variations of these parameters. For fast incident electrons both γ and $\epsilon_{\min}/E$ are very small numbers, generally being on the order of 10^{-3}. In the limit as γ and $\epsilon_{\min}$ tend to zero both the elastic and inelastic cross sections diverge, but one can easily verify that they have the symbolic expansions

$$\sigma(E,\mathbf{n}\cdot\mathbf{n}') = \sigma_T(E)\delta(\mathbf{n}-\mathbf{n}') - e^4Z^2(2E)^{-2} \times\ln(1/\gamma)\delta'(1-\mathbf{n}\cdot\mathbf{n}')+O(1), \quad (12)$$

$$\omega(E,E') = \omega_T(E)\delta(E-E') - 2\pi Ze^4E^{-1} \times\ln(2E/I)\delta'(E-E')+O(1), \quad (13)$$

[14] R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill Book Company, Inc., New York, 1960).

[15] L. Landau, J. Phys. 8, 201 (1944).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where the symbol $O(1)$ has been used to denote finite terms independent of γ or ϵ_{min} plus terms which vanish as γ or ϵ_{min} tends to zero and we have used Eq. (6) to simplify (13). Referring to Eq. (3) we see that the terms proportional to σ_T and ω_T drop out so the dependence of R on ϵ_{min} and γ is only logarithmic.

Upon inserting the cross sections (4) and (5) into Eq. (2), one can verify by a simple change of variables that if $R(\mathbf{n},E;\mathbf{n}',E')$ is a solution so is $\lambda R(\mathbf{n},\lambda E;\mathbf{n}',\lambda E')$ where λ is any positive number, *provided* that one neglects the energy dependence of the logarithms $\ln(1/\gamma)$ and $\ln(2E/I)$ appearing in (12) and (13). When the latter is a good approximation we can apply our uniqueness "theorem" which then yields

$$R(\mathbf{n},E;\mathbf{n}'E')\approx\lambda R(\mathbf{n},\lambda E;\mathbf{n}',\lambda E'). \quad (14)$$

Physically this means that increasing the incoming and outgoing energies in the same proportion will leave the whole backscattering process unchanged, except for a scaling factor which is precisely that needed to keep the integrated coefficient at a constant value.

While it is reassuring to derive this well verified symmetry from theory, it is perhaps more important to investigate the conditions under which it is valid. To begin with, Eq. (14) does not hold unless we can neglect the energy dependence of the logarithms in (12) and (13). Roughly speaking, this requires that the incident energy be greater than the K-shell binding energy of the scatterer. For carbon the K-shell energy is roughly 300 eV and one finds experimentally that the total backscattering coefficient is very nearly constant for $E\gtrsim 500$ eV.[2,3] On the other hand, the K-shell energy of silver is about 25 keV and one indeed finds that the number of backscattered electrons does not become constant until E is on the order of 15 keV.[2,3] A second breakdown of (14) occurs when the angle of incidence is near 90° (Fig. 1). In this case the incoming electrons can take full advantage of the forward peak in the Rutherford cross section which is extremely sensitive to γ. With an incident angle of 80°, Kanter[4] has found that the energy distribution of electrons backscattered from aluminum varies considerably for incident energies between 10 and 70 keV even though the K-shell energy of aluminum is on the order of 1.5 keV. Since backscattering is largest at grazing incidence, this breakdown of (14) for $\theta\approx 90°$ will have a disproportionate effect when the incident flux is isotropic. Formally, Eq. (14) breaks down near $\theta=90°$ because $\ln\gamma$ multiplies an angular derivative which becomes very large when the initial direction is parallel to the surface of the target. Finally, if the relativistic cross sections, rather than those given by (4) and (5) are used in Eq. (2), one finds that the resulting equation has a different dependence on the energy variables and (14) no longer holds. Experimentally[5] one finds that in the relativistic region the backscattering coefficient is not independent of the incident energy, but shows a marked decrease from the nearly constant nonrelativistic value.

4. APPROXIMATIONS

Most theories of electron transport are based on the assumption that electrons lose energy continuously according to the usual range-energy relationship. This is definitely an approximation since energy loss, like angular deflection, is a stochastic process. To determine the accuracy of this approximation when applied to backscattering, we must examine the integrals on the left-hand side of Eq. (3). For simplicity, we will consider only the first integral and assume that the integration over the angular delta function has already been carried out. The inelastic cross section $\omega(E,E'')$ is very strongly peaked near $E=E''$, which suggests that we change the lower limit of integration from E' to zero and, keeping E' fixed, expand $R(\mathbf{n},E'';\mathbf{n}',E')$ in a Taylor series around $E''=E$. The zeroth-order term in the expansion is exactly cancelled by the $\delta(E-E'')\omega_T(E)$ term and the first integral becomes

$$\sec\theta\left[\int_0^E \omega(E,E'')(E-E'')dE''\right]\frac{\partial}{\partial E}R(\mathbf{n},E;n',E')$$

$$-\tfrac{1}{2}\sec\theta\left[\int_0^E \omega(E,E'')(E-E'')^2dE''\right]\frac{\partial^2}{\partial^2 E}$$

$$\times R(\mathbf{n},E;\mathbf{n}',E')+\cdots. \quad (15)$$

According to (6) the first term in (15) can be written as

$$-\sec\theta\frac{dE}{dx}\frac{\partial}{\partial E}R(\mathbf{n},E;\mathbf{n}',E').$$

On the other hand, in the continuous approximation the probability of energy loss per unit path length $\omega(E;E')$ would be

$$(dx)^{-1}\delta[E-E'-(dE/dx)dx].$$

In this case ω_T is simply $(dx)^{-1}$, and upon inserting these expressions into the first term on the left-hand side of (3) and taking the limit $dx\rightarrow 0$, one finds that the continuous energy loss approximation is exactly equivalent to keeping only the first term in (15). Since $\omega(E,E'')$ is peaked near $E=E''$, the coefficient of the first term in (15) will be considerably larger than the coefficients of the higher order terms so that dropping these terms will be a good approximation unless R is so rapidly varying that the higher derivatives are large. To estimate these derivatives one must remember that we are considering R as a function of E for fixed E'. Experimentally one finds that for low-Z targets, R is a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

slowly varying function and the approximation should be very good but for high-Z targets R changes by an order of magnitude between $E/E'=1$ and $E/E'=2$. A rough calculation indicates that for Z greater than about 40, the second term in (15) can be nearly as large as the first. One cannot, therefore, expect a theory based on a continuous energy loss to give detailed, quantitative results for high-Z materials. This conclusion is in agreement with Monte Carlo calculations carried out by MacCallum,[16] who found that calculations based on the usual range-energy relationship gave the details of the energy distribution of backscattered electrons correctly only for low-Z scatterers.

The easiest way to incorporate this approximation into the present formalism is through Eq. (9). If the energy loss is assumed to take place continuously, the function $P(E,E';t\sec\theta)$ defined in (10) is obviously just a delta function whose argument is determined by the range-energy relation which in terms of t is given by $dE/dt=-E^{-1}\ln(2E/I)$. If we want our solution to obey Eq. (14) we must neglect the energy variation of the logarithm in which case the range-energy relation can be integrated analytically and we find

$$P(E,E';t\sec\theta)=2E'\delta(E^2-E'^2-Ct\sec\theta), \quad (16)$$

where $C=2\ln(2E/I)$. There is, of course, some ambiguity in assigning a numerical value to the "constant" $\ln(2E/I)$. However, one will note that (16) is simply a restatement of the Thompson-Whiddington law[9] and Terrill[17] has found that this simplification of the range-energy relationship is fairly accurate provided that one takes $C\approx 10$. Since one might argue that this value of C is not the proper one to use in backscattering problems, it is interesting to observe that C can be determined directly by studying the behavior of R in the limit $E'\to E$. Clearly an electron which has lost very little energy cannot have spent enough time in the target to have been scattered more than once, which means that in the limit $E'\to E$ the first-order term in an expansion of R in powers of Z is exact. This term can be obtained simply by setting $R=0$ on the right hand side of Eq. (9).

At this point it is convenient to introduce some notation which will be used throughout the rest of this section. Since we will be interested only in solutions which satisfy Eq. (14), it is advantageous to work with the dimensionless function

$$r(\mathbf{n},\mathbf{n}';\xi)=E\cos\theta' R(\mathbf{n},E;\mathbf{n}',E'), \quad (17)$$

where $\xi=E'/E$ and the factor $\cos\theta'$ has been added to facilitate comparison with experimental data. That r is a function only of E'/E follows directly from (14).

Substituting (16) into Eq. (9) and integrating over the delta functions, one easily verifies that to order Z

$$r(\mathbf{n},\mathbf{n}';\xi)=\frac{Z}{4\pi C}\frac{\xi\cos\theta\cos\theta'}{\cos\theta+\xi^2\cos\theta'}\frac{1}{(1+\mathbf{n}\cdot\mathbf{n}'+\gamma)^2}. \quad (18)$$

Taking $C=10$, we find that for $\xi=1$ and $\theta=0$ (normal incidence) r should behave as shown in Fig. 3. The curves are in excellent agreement with the experimental points of Kulenkampff and Rüttiger.[6] One will note that in the limit $\xi\to 1$ it is not an approximation to neglect the energy dependence of the $\ln(2E/I)$ term in the range-energy relation which means that for $\xi=1$ Eq. (18) is, within the approximation of continuous energy loss, exact and C has the unambiguous value of $2\ln(2E/I)$ where E is the incident energy. Since the experimental points were taken at 30 keV, this would yield an ionization potential $I\approx 400$ eV independent of Z whereas one usually takes $I\approx(13.4\text{ eV})Z$. This is indeed a remarkable occurrence for which there appears to be no immediate explanation.

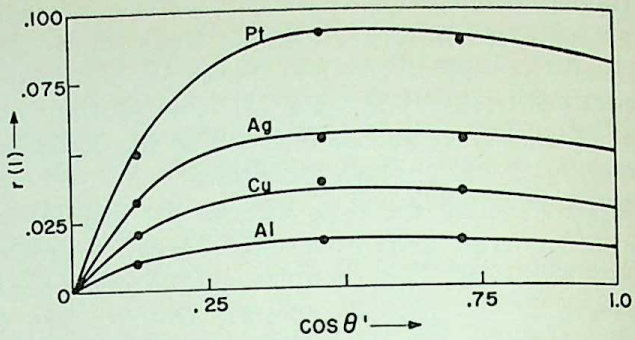


FIG. 3. Behavior of r for $\xi=1$ and normal incidence. The dots are the experimental points of Kulenkampff and Rüttiger (Ref. 6).

Recently there have been attempts to explain backscattering in terms of a single large-angle scattering.[9,10] Since a theory that includes only a single scattering necessarily corresponds to the first term in an expansion of R in powers of Z, it must be exact in the limit $Z\to 0$, and one might hope that it would be valid for low-Z materials of experimental interest. Further motivation for this approach can be found in the fact that for Z between 6 and 40 the total backscattering coefficient for normal incidence is about $0.013Z$ which is suggestive of a first-order term in Z. However, for $\theta=0$, Eq. (18) yields a total coefficient of $[8C]^{-1}[1-\ln 2]Z+0(Z^2)$ which for $C=10$ is about $0.004Z$. Everhart[9] has suggested treating C as a phenomenological parameter which should be adjusted to give the observed backscattering. In view of the accuracy with which Terrill's value reproduces the experimental points at $\xi=1$ this is hardly a permissible procedure. Moreover, the energy distribution of backscattered electrons predicted by (17) is in violent disagreement with experiment so we are forced to the conclusion that a single-scattering theory is not valid for realistic values of Z. Apparently the linear behavior of the total coefficient for low-Z materials is accidental.

[16] C. MacCallum, Bull. Am. Phys. Soc. 5, 379 (1960).
[17] H. Terrill, Phys. Rev. 22, 101 (1923).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Having found that a theory based on a single large-angle scattering is not realistic, one might be tempted to go to the opposite extreme and use a small-angle approximation in which one keeps only the forward peak of the cross section. The possibility of constricting a quantative theory of this sort is imbedded in what is perhaps the most interesting question in backscattering theory: Which is the more important, small- or large-angle scattering? One way to answer this question is simply to chop off the small-angle part of the Rutherford cross section, calculate the resulting backscattering coefficient, and compare the results with experiment. We have made a number of such calculations by iterating Eq. (9) in the continuous energy-loss approximation (16). Other motivations for the calculations were the hope that we might obtain quantitative results for low-Z targets and to find out if matrix equations like (9) are suitable for machine calculations. More specifically, we replaced the Rutherford cross section $[1-\cos\Theta+\gamma]^{-2}$ with $[1-\cos\Theta]^{-2}\exp[-b^2/(1-\cos\Theta)^2]$ and iterated (9) starting with $R=0$. As long as Z was not large or b too small, the iterates converged satisfactorily becoming almost constant after about six iterations. However, for large Z or very small b satisfactory convergence could be obtained only for values of ξ close to one. This was due to the fact that the matrix S_F is the difference between two terms which partially cancel each other, and if either S_F or R becomes too large or rapidly varying, the equation becomes numerically unstable. For this reason, the method is not practical for applications of the theory and it would not be worth while to go into the details.

Some typical results obtained for $b=0.1$ are shown in Figs. 4 and 5. While the calculated curves lie considerably lower than the experimental ones, the results are remarkable considering that we have in effect said that if an electron is not scattered through an angle greater than about 25°, it is not scattered at all. Particularly surprising is the fact that the calculated curve for copper is about as close to the experimental one as it is for aluminum, even though the atomic number of copper is over twice that of aluminum. This appeared to be a general trend. Our results for carbon and silver were percentage wise, in about the same agreement as those for copper and aluminum. Apparently the physical reason for this is that the increased probability of small angle scattering in the higher Z materials is being compensated for by the shorter time which a typical backscattered electron spends in the target. One will also note that the relative discrepancy between the calculated and observed curves is larger for $\theta'=83°$ than for $\theta'=43°$. This is to be expected since the points at $\theta'=83°$ corresponds to deflection through a smaller angle, which is more likely to be effected by small-angle scattering. As the cutoff angle was decreased, the calculated points rose towards the experimental values. However, the difference between the curves computed with cutoffs of 25° and 12° was only about 15% which indicates that the bulk of the discrepancy is due to scattering at angles well inside the forward peak of the Rutherford cross section. The conclusion is that quite independently of Z, about half the backscattering arises from a relatively small number of large-angle scatterings, with the remainder being due to the diffusive effects of small-angle scattering. In retrospect, this is not a surprising result since the mean-square scattering angle is on the order of γ so one needs roughly γ small-angle scatterings to deflect an electron through 90°, but about one out of every γ collisions will give rise to a single scattering through an angle greater than 90°.

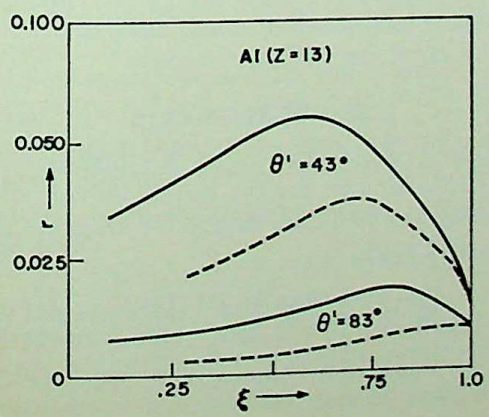


FIG. 4. Results of the cutoff calculations (dashed line) for normal incidence compared with the experimental data (Ref. 6) of Kulenkampff and Rüttiger (solid line).

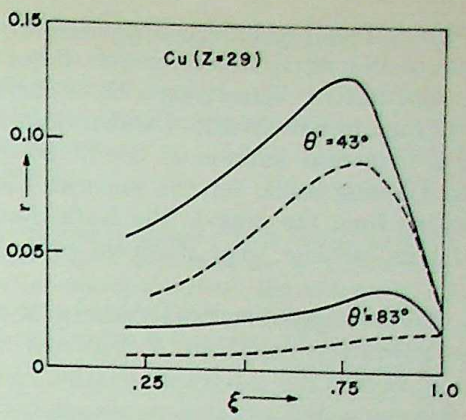


FIG. 5. Results of the cutoff calculations (dashed line) for normal incidence compared with the experimental data (Ref. 6) of Kulenkampff and Rüttiger (solid line).

5. A POSSIBLE METHOD FOR NUMERICAL SOLUTION

The main result of the previous section was that backscattering is a very complicated problem. For one thing, we found that both the forward peak of the scattering cross section and the details of its large-angle tail are important. Also, it appears that for high-Z materials one must include the stochastic nature of energy loss. For these reasons it seems very doubtful that a simplified treatment will ever lead to more than

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

a crude estimate of the total coefficient. Barring the appearance of an entirely new formalism, if one wishes to make quantitative calculations, he is faced with either resorting to the Monte Carlo method or attempting a numerical solution of one of the present equations. Leaving aside, for the moment, the difficulties arising from the peak in the Rutherford cross section, let us consider some pertinent properties of Eq. (2).

The factors $\sec\theta$ appearing in (2) blow up at $\theta=90°$. Since the laboratory coefficient $R\sec\theta\cos\theta'$ is well behaved at $\theta=90°$ the integrand remains finite. In carrying out our cutoff calculations we approximated the angular integrals with a Gaussian quadrature scheme which avoids the ambiguous points at $\theta=90°$ and found that these apparent divergences caused no difficulty.

Since the problem has rotational symmetry around an axis perpendicular to the target, $R(\mathbf{n},E;\mathbf{n}',E')$ has the property that its average over the azimuthal angle of either $\mathbf{n}$ or $\mathbf{n}'$ is equal to its average over the azimuthal angle of both $\mathbf{n}$ and $\mathbf{n}'$. Using this property, one readily verifies that averaging (2) over the azimuthal angles of $\mathbf{n}$ and $\mathbf{n}'$ yields an equation for the corresponding average of R. The resulting equation is still of the matrix form (8) and the same techniques can be applied to both the full and averaged equations. This procedure will greatly reduce the number of points needed to carry out the angular integrations.

Another simplification arises if the solution is required to satisfy (14). As one would expect, (2) then becomes an equation for the dimensionless function r defined in (17) and the number of energy variables is cut in half. We have pointed out that $R(\mathbf{n},E;\mathbf{n}',E')$ is independent of reflection coefficients whose energy indices are greater than E or less than E' which means that the integral equation for r is of the Volterra type as far as the ξ variable is concerned. One can therefore compute r out to some minimum value of ξ without having to find the function for smaller values of ξ. The validity of forcing R to satisfy (14) is, however, not *a priori* known.

The major stumbling block will, of course, be the strong forward peak in the Rutherford cross section. One will note, however, that the matrix W is also ill behaved, but when it was exponentiated to obtain Eq. (9) its singular behavior ceased to cause any difficulty. The latter equation can be obtained by multiplying (8) from both right and left by $\exp(-Wt)$, noticing that the left-hand side is a total derivative and integrating from zero to infinity. More generally, suppose we take any matrix function $A(t)$ which has the properties, $A(0)=I$ and $A(\infty)=0$, and carry out the same manipulations. One easily verifies that the analog of (9) is then

$$R=\frac{Z}{8\pi}\int_0^\infty A(t)[S_B+B_1(t)R+RB_2(t)+RS_BR]A(t)dt,$$

$$B_1(t)=\frac{8\pi}{Z}\left[-A^{-1}(t)\frac{d}{dt}A(t)-W+\left(\frac{Z}{8\pi}\right)S_F\right], \tag{19}$$

$$B_2(t)=\frac{8\pi}{Z}\left[-\left[\frac{d}{dt}A(t)\right]A^{-1}(t)-W+\left(\frac{Z}{8\pi}\right)S_F\right].$$

The game is now to choose A so that the matrices B_1 and B_2 are well behaved. The ideal choice would be $\exp[-(W-Z/8\pi S_F)t]$, but from the analog of Eq. (11) one finds that the $(\mathbf{n},E;\mathbf{n}',E')$th element of this matrix is the probability that an electron will be scattered from the state $(\mathbf{n},E)$ to the state (n',E') while passing through a slab of thickness t, provided that one only includes scatterings such that the electron is always traveling into the target. Evaluating this matrix would be as difficult a problem as solving the original equation. On the other hand, the theory of small-angle multiple scattering[18,19] has provided a number of approximate evaluations of this matrix. These approximations are only valid for small-angle deflections, but one will note that this is all that is needed to insure that B_1 and B_2 are well behaved. The scheme is, therefore, to calculate $\exp[-(W-Z/8\pi S_F)t]$ in some tractable small-angle approximation, use this matrix for $A(t)$ in (19), and solve the resulting equation by iteration starting with $R=0$. Any reasonable approximation for the above exponential should remove most of the small-angle scattering from B_1 and B_2, and the successive iterates should converge like an expansion in the number of large-angle scatterings. Our experience with the cutoff calculations described in the previous section indicate that such a series would converge rapidly after about five iterations. Furthermore, any iterative solution to (19) is clearly well behaved for small Z, so if the iterates converge they are guaranteed to yield the correct solution. The author has not investigated the above scheme in detail but preliminary investigations indicate that it may lead to a practical method for obtaining numerical results.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Dr. C. J. MacCallum for his helpful advice and encouragement and to K. G. Adams for his assistance in the numerical work.

[18] W. Scott, Rev. Mod. Phys. **35**, 231 (1963).
[19] H. Bethe, M. Rose, and L. Smith, Proc. Am. Phil. Soc. **78**, 573 (1938).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Growing Helical Density Waves in Semiconductor Plasmas*

C. E. HURWITZ AND A. L. MCWHORTER
Lincoln Laboratory† and Electrical Engineering Department, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 2 December 1963)

An experimental and theoretical investigation is made of growing screw-shaped plasma density waves in a semiconductor bar subjected to parallel electric and magnetic fields. A particularly simple mode is used, requiring only a thermal-equilibrium electron-hole plasma, low-recombination surfaces, and moderate fields. It is shown that in extrinsic material the growth is spatial, corresponding to stable traveling-wave amplification, while for nearly equal densities of positive and negative carriers the wave is absolutely unstable and corresponds to the oscillistor phenomenon. Experimental observations of the waves were made in germanium at and above room temperature for frequencies from 20–400 kc and with electric and magnetic fields from 25–60 V/cm and 0–11 kG, respectively. The growth rates and phase characteristics were found to be in excellent agreement with theory and gain in excess of 35 dB/cm was obtained. At higher temperatures, corresponding to nearly intrinsic material, evidence of instability was found in accordance with the theoretical prediction.

I. INTRODUCTION

HELICAL plasma density waves were originally proposed by Kadomtsev and Nedospasov[1] to account for an instability that occurs in the positive column of a gas discharge in a longitudinal magnetic field. Subsequently, Glicksman[2] suggested that a slightly generalized form of this theory could explain the oscillistor effect in semiconductors.[3–5] In its simplest form the oscillistor is a semiconductor bar that exhibits terminal voltage and/or current oscillations when an electron-hole plasma is created by injection or light generation and sufficiently large parallel electric and magnetic fields are applied. Further refinements of the theory for semiconductors and experimental confirmation of the helical nature of the oscillistor instability have been provided by several other authors,[6–13] leaving no doubt as to the correctness of Glicksman's basic idea. However, the feature from which the oscillistor takes its name, the terminal voltage-current oscillation, has not been adequately accounted for up to now, nor has it been determined whether the oscillation is caused by an absolute instability or a convective instability with feedback.[14–19]

The present work indicates that the oscillistor instability is an absolute one and that the terminal oscillations can result from nonlinearities alone. The bulk of the paper, however, is devoted to the theoretical and experimental investigation of a completely different aspect of these helical waves, that of traveling-wave amplification. In this study, a preliminary account of which has been given earlier,[20] a mode is utilized that is much simpler than the one considered by previous authors and which involves only the thermal-equilibrium hole-electron plasma. In these other treatments the growth was always assumed to be temporal, representing instability, and the possibility of stable amplification was not considered, perhaps because the high degree to which the growth and propagation constants are influenced by the net difference in hole and electron densities was not fully appreciated. It will be shown that using nothing more than a uniform bar of germanium at room temperature with a low-recombination surface, Ohmic contacts, and moderate applied electric and magnetic fields, controlled amplification of as much as 35 dB/cm can be achieved. In the limit of nearly equal hole and electron densities (intrinsic material or material with a high level of injected carriers) the helical wave becomes absolutely unstable, corresponding to the oscillistor situation mentioned above.

---

* Based on a thesis submitted by C. E. Hurwitz to the Massachusetts Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

† Operated with support from the U. S. Air Force.

[1] B. B. Kadomtsev and A. V. Nedospasov, J. Nucl. Energy 1, 230 (1960).

[2] M. Glicksman, Phys. Rev. 124, 1655 (1961).

[3] I. L. Ivanov and S. M. Ryvkin, Zh. Techn. Fiz. 28, 774 (1958) [English transl.: Soviet Phys.—Tech. Phys. 3, 722 (1958)].

[4] J. Bok and R. Veilex, Compt. Rend. 248, 2300 (1958).

[5] R. D. Larrabee and M. C. Steele, J. Appl. Phys. 31, 1519 (1960).

[6] F. Okamoto, T. Koike, and S. Tosima, J. Phys. Soc. Japan 17, 804 (1962).

[7] Y. H. Ichikawa, S. Misawa, and Y. Sasakura, Progr. Theoret. Phys. (Kyoto) 27, 1277 (1962).

[8] T. Misawa, Japan. J. Appl. Phys. 1, 67 (1962).

[9] T. Misawa, Japan. J. Appl. Phys. 1, 131 (1962).

[10] B. Ancker-Johnson, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter 1962* (The Institute of Physics and the Physical Society, London, 1962), p. 141; Bull. Am. Phys. Soc. 7, 496 (1962).

[11] Ø. Holter, Phys. Rev. 129, 2548 (1963).

[12] R. D. Larrabee, J. Appl. Phys. 34, 880 (1963).

[13] T. Misawa and T. Yamada, Japan. J. Appl. Phys. 2, 19 (1963).

[14] R. Q. Twiss, Proc. Phys. Soc. (London) B64, 654 (1951).

[15] P. A. Sturrock, Phys. Rev. 112, 1488 (1958).

[16] L. D. Landau and E. M. Lifshitz, Fluid Mechanics (Pergamon Press, Ltd., London, 1959), pp. 111–114.

[17] Y. B. Fainberg, V. I. Kurilko, and V. D. Shapiro, Zh. Techn. Fiz. 31, 633 (1961) [English transl.: Soviet Phys.—Tech. Phys. 6, 459 (1961)].

[18] R. V. Polovin, Zh. Techn. Fiz. 31, 1220 (1961) [English transl.: Soviet Phys.—Tech. Phys. 6, 889 (1962)].

[19] A. Bers and R. Briggs, Research Laboratory of Electronics, MIT, Quarterly Progress Report No. 71, 1963, p. 122 (unpublished).

[20] C. E. Hurwitz and A. L. McWhorter, Phys. Rev. Letters 10, 20 (1963); Bull. Am. Phys. Soc. 8, 206 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

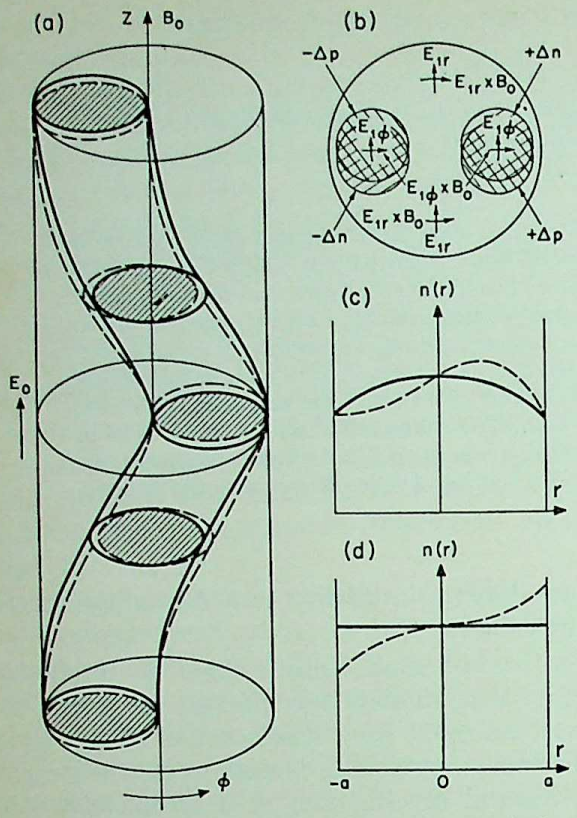


FIG. 1. Right-handed helical perturbation for $|m|=1$. (a) Perturbed-electron distribution indicated by dashed lines and perturbed hole distribution by solid lines. [After Hoh and Lehnert (Ref. 21).] (b) Cross-sectional view showing charge separation and resulting perturbed electric fields E_1. Directions of carrier flow due to the combined effect of E_1 and B_0 are indicated by the vectors $\mathbf{E}_1 \times \mathbf{B}_0$. (c) and (d) Carrier density versus radius along a horizontal diameter through the cross section of (b) for the equilibrium-gradient mode and the surface-density mode, respectively. Solid lines correspond to unperturbed distributions and dashed lines to distributions altered by radial and azimuthal flows.

II. PHYSICAL MODEL

A physical model of the growth mechanism of the helical wave was developed by Hoh and Lehnert[21,22] for the gas-discharge case. We will include an additional part of the mechanism and modify the picture to apply to semiconductors. The argument is begun by assuming a small, quasineutral, screw-shaped perturbation of positive and negative carriers to be superimposed on the steady-state unperturbed distribution. The general form, in cylindrical coordinates, would be $F(r)\exp(i\omega t - ikz - im\phi)$ with $F(r)$ determined by the boundary condition at the wall; however, for the present argument we consider only the lowest mode $m=\pm 1$, as the full mathematical treatment shows that this mode has the lowest threshold for growth. The perturbation is shown schematically in Fig. 1(a) for m and k having opposite signs (a right-handed helix). The longitudinal electric field E_0 tends to separate the superimposed positive and negative screws axially, which is equivalent to a rotation of one screw relative to the other, as shown in Fig. 1(a) and in the cross-sectional view of Fig. 1(b). The resulting charge separation creates both an azimuthal and a radial electric field, $E_{1\phi}$ and E_{1r}, respectively. The radial field was neglected by Hoh and Lehnert but it is no less important than the azimuthal field in the description of the total growth mechanism. In the linear approximation these fields, together with the longitudinal magnetic field B_0, act on the unperturbed distribution to produce a radial and azimuthal flow of particles, indicated in Fig. 1(b) by the vectors $\mathbf{E}_{1\phi} \times \mathbf{B}_0$ and $\mathbf{E}_{1r} \times \mathbf{B}_0$. If the unperturbed distribution has a radial gradient, as it does in a gas discharge or in the usual form of the oscillistor (due to the finite recombination of injected carriers at the walls), this flow is not divergenceless and hence can feed particles from the main distribution into the screw with the proper phase, thus producing a growth of the perturbation. This may be seen more clearly in Fig. 1(c), where the carrier density is plotted as a function of radius along a diameter horizontally through the cross section of Fig. 1(b). For the signs of E_0 and B_0 chosen, the flow adds carriers to the assumed excess in the region $r>0$ and removes carriers from the assumed deficit in the region $r<0$, thus enhancing the perturbation.

When the influx of particles is sufficiently large to overcome the dissipative effects of diffusion and recombination the helix will grow. It may thus be expected that for a given value of electric field E_0 there will be a threshold value of magnetic field B_{0c} above which growth is observed. There should also be an optimum wavelength, for if the wavelength is too long the charge separation and therefore the driving electric fields will be too small, while for a very short wavelength the diffusive effects will take over and dissipate the perturbation. If the sample or plasma column is long enough that end effects are negligible, the wavelength should be independent of the sample length. For the electric and magnetic fields as shown in Fig. 1(a) the screw of opposite sense will clearly be attenuated since the radial and azimuthal flows now reverse directions and tend to destroy the perturbation. In general it may be stated[9] that the helix which can exhibit growth is right- (left-) handed for the electric and magnetic fields parallel (antiparallel).

In a semiconductor a radial gradient of the unperturbed distribution can be produced in thermal equilibrium by proper doping (e.g., by indiffusion or outdiffusion of impurities), eliminating the need for a steady-state injected or optically-generated plasma with all of its inherent complications. Of course, the radial gradients of the holes and electrons in this "equilibrium-gradient mode" will be in opposite directions, producing a partial cancellation in the growth process, but as shown in a calculation of the threshold for this mode in Appendix A there can still be a net buildup of the perturbation.

However, except for this calculation, we will be concerned with an even simpler mode of semiconductor operation requiring no unperturbed radial gradient at

[21] F. C. Hoh and B. Lehnert, Phys. Rev. Letters 7, 75 (1961).
[22] F. C. Hoh, Phys. Fluids 5, 22 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

all, merely a sufficiently low surface recombination. In this mode, called here the "surface-density mode," growth occurs because the radial and azimuthal flows can pile up or deplete the carriers at the surface in the proper phase, as illustrated in Fig. 1(d). The role of the dc-electric and magnetic fields in producing the flow is the same as before. Experimentally the surface-density mode is easier to achieve than the equilibrium-gradient mode, and also the mathematical theory for the surface-density mode can be formulated with fewer approximations.

In addition to the growth properties, the motion of the wave can also be ascertained from the physical model. As is well known, a quasineutral density perturbation under the influence of an applied electric field E_0 propagates in the direction of minority carrier drift with a velocity $v_d=\mu_a E_0$, with μ_a, the ambipolar drift mobility, being given by

$$\mu_a=[\mu_e\mu_h(n_0-p_0)]/(n_0\mu_e+p_0\mu_h),$$

where n_0 and p_0 are the equilibrium densities of electrons and holes, and μ_e and μ_h are their respective mobilities. Thus, the helix performs a bodily translation with the ambipolar drift velocity v_d. In addition, there is a rotation of the helix due to diffusion and conduction of carriers across the magnetic field. (Another mechanism for rotation, due to the radial ambipolar electric field, also exists when the unperturbed carrier distribution has a radial gradient.) The total motion of the helix is composed of these two parts. For extrinsic semiconductors the translation will dominate, while for intrinsic semiconductors the ambipolar mobility is zero and the motion is therefore purely rotational. Discussion of the direction of rotation will be postponed until the next section where the necessary mathematical expressions are derived. It should be noted that because of the nature of a helix both motions when viewed in any cross-sectional plane produce a rotation of the carrier density and field patterns about the axis.[23]

The ambipolar drift of the helix also plays a crucial role in distinguishing instability from stable growth. A detailed discussion is presented in the next section, but to complete the physical picture the basic ideas are outlined here. Three fundamental processes govern the change in amplitude of the perturbation with time at any point in space: the growth mechanism previously described, diffusion, and ambipolar drift down the bar. Above threshold the growth is enough to overcome diffusion, but for a sufficiently large drift velocity the disturbance is swept away, growing as it moves. At any point in space the perturbation remains bounded in time, and the result is simply a spatially growing wave. However, if the ambipolar drift velocity is very small the perturbation may build up faster than it can diffuse and drift away. At a point in space it may then grow temporally, corresponding to an absolute instability. Clearly the first case is pertinent to extrinsic semiconductors, and the latter to semiconductors with nearly equal electron and hole densities and to gaseous plasmas.

It should also be mentioned that Gurevich and Ioffe[24] have proposed a very similar mechanism which can lead to growing density waves in thin semiconductor slabs. In their model a steady-state density gradient and electric fields are set up either by illumination of one surface or by the Hall effect. Then the addition of a wave-like density perturbation creates a perturbed electric field through charge separation in the steady-state electric and magnetic fields. As in the case of the helical wave, this perturbed field acts on the unperturbed distribution to produce a carrier flow in the correct direction to enhance the perturbation. The applied electric and magnetic fields perform different functions in the chain of physical argument, but the fundamental process of coupling an unperturbed density gradient with a perturbed electric field to achieve growth is basically the same as for the helical wave. Experimental observations of instability under conditions similar to those considered by Gurevich and Ioffe have been reported[25,26] and may well be explained in terms of their model.

III. MATHEMATICAL THEORY

With the preceding physical picture in mind, we now formulate a quantitative mathematical description of the helical wave in the surface-density mode. The equilibrium-gradient mode is considered in Appendix A. Using the equations of motion and continuity for electrons and holes, the dispersion relation governing the growth and propagation of the helical wave will be derived. From this equation will then follow both the threshold conditions and the growth and phase constants for stable growth as well as the threshold for instability.

Consider a cylindrical bar of semiconductor, long in the z direction so that end effects may be neglected. The bar is assumed to be uniform and in thermal equilibrium and therefore to contain a uniform density n_0 of electrons and p_0 of holes with $n_0-p_0=N_D-N_A$, where N_D and N_A are the densities of ionized donor and acceptor impurities, respectively. Under the assumption that the times and distance of interest are much larger than the mean free times and paths of the free carriers, we may write the following equations to describe the behavior of

[23] The contribution of the drift motion to this rotation was first pointed out by Okamoto *et al.* (Ref. 6).

[24] L. E. Gurevich and I. V. Ioffe, Fiz. Tverd. Tela 4, 2641, 2964 (1962) [English transl.: Soviet Phys.—Solid State 4, 1938, 2173 (1963)].

[25] S. Nakashima and Y. Miyai, J. Phys. Soc. Japan 18, 1219 (1963).

[26] S. Nakashima and Y. Noguchi, Japan. J. Appl. Phys. 2, 307 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the carriers under applied electric and magnetic fields:

$$\mathbf{J}_e = nq\mu_e\mathbf{E} + qD_e\nabla n - \mu_e\mathbf{J}_e\times\mathbf{B}, \quad (1)$$

$$\mathbf{J}_h = pq\mu_h\mathbf{E} - qD_h\nabla p + \mu_h\mathbf{J}_h\times\mathbf{B}, \quad (2)$$

$$\partial n/\partial t = (1/q)\nabla\cdot\mathbf{J}_e + \gamma, \quad (3)$$

$$\partial p/\partial t = -(1/q)\nabla\cdot\mathbf{J}_h + \gamma, \quad (4)$$

$$\nabla\cdot\mathbf{E} = -q(n-p)/\epsilon. \quad (5)$$

The subscripts e and h refer to the electrons of density n and holes of density p, respectively; μ is the mobility, D is the diffusion constant, γ is the net rate of bulk generation (which in the absence of trapping is the same for both carriers), ϵ is the dielectric constant of the semiconductor, q is the magnitude of the electronic charge, J is the current density, and E and B are the electric and magnetic fields, respectively.

In order to effect a considerable simplification in the mathematics it will be assumed throughout that the square of the Hall angle for each carrier is small; i.e., that $\mu_{e,h}{}^2B^2\ll 1$. Under this assumption and also assuming that the self-magnetic fields are negligible with respect to the applied uniform magnetic field B_0, Eqs. (1)–(4) may be combined to yield

$$\partial n/\partial t = n\mu_e\nabla\cdot\mathbf{E} + \mu_e\mathbf{E}\cdot\nabla n + D_e\nabla^2 n + \mu_e{}^2\mathbf{E}\cdot(\nabla n\times\mathbf{B}_0), \quad (6)$$

and

$$\partial p/\partial t = -p\mu_h\nabla\cdot\mathbf{E} - \mu_h\mathbf{E}\cdot\nabla p + D_h\nabla^2 p + \mu_h{}^2\mathbf{E}\cdot(\nabla p\times\mathbf{B}_0). \quad (7)$$

The bulk generation term has been omitted as it is a simple matter to reinsert it if desired near the end of the calculation.

Equations (6) and (7) are now linearized for small perturbations about equilibrium by writing

$$n = n_0 + n_1, \quad p = p_0 + p_1, \quad E = E_0 - \nabla\psi_1, \quad (8)$$

where the subscripts 0 and 1 refer to the equilibrium quantities and their perturbations, respectively. Assuming quasineutrality (i.e., $n_1\approx p_1$, but $\nabla^2\psi_1\neq 0$), eliminating ψ_1 from the linearized equations, and restricting ourselves to the particular case of interest where E_0 and B_0 are both in the z direction, we have the familiar ambipolar continuity equation for n_1:

$$\partial n_1/\partial t = D_a\nabla^2 n_1 - \mu_a E_0(\partial n_1/\partial z), \quad (9)$$

where

$$\mu_a \equiv [\mu_e\mu_h(n_0-p_0)]/(n_0\mu_e + p_0\mu_h)$$

and (10)

$$D_a \equiv (n_0\mu_e D_h + p_0\mu_h D_e)/(n_0\mu_e + p_0\mu_h)$$

are the ambipolar mobility and diffusion constant, respectively.

To obtain the normal modes for the problem, the perturbations n_1 and ψ_1 are Fourier analyzed in cylindrical coordinates (r,ϕ,z) in the form

$$n_1 = N_1(r)\exp(i\omega t - ikz - im\phi),$$
$$\psi_1 = \Psi_1(r)\exp(i\omega t - ikz - im\phi), \quad (11)$$

and (9) is solved for $N_1(r)$, the radial part of n_1, with the result

$$N_1(r) = c_1 I_m(\beta r), \quad (12)$$

where $I_m(\beta r)$ is the modified Bessel function of the first kind of order m, c_1 is an arbitrary constant, and

$$\beta^2 \equiv [(\omega - \mu_a E_0 k)/D_a] + k^2. \quad (13)$$

The potential $\Psi_1(r)$ is obtained from the continuity equations (6) and (7), which after linearization and for E_0 and B_0 in the z direction may be combined in the form

$$\nabla^2\psi_1 + A\nabla^2 n_1 = 0, \quad (14)$$

where

$$A \equiv -i\frac{E_0 k(\mu_e+\mu_h)}{(n_0\mu_e+p_0\mu_h)(\beta^2-k^2)} + \frac{D_e-D_h}{n_0\mu_e+p_0\mu_h}. \quad (15)$$

Therefore,

$$\psi_1 = An_1 + \text{a solution of Laplace's equation.} \quad (16)$$

The appropriate solution of Laplace's equation is a modified Bessel function $I_m(kr)$, and as a result

$$\Psi_1(r) = Ac_1 I_m(\beta r) + c_2 I_m(kr). \quad (17)$$

To complete the solution, the boundary condition at the surface $r=a$ of the cylinder must be applied. These conditions have the form

$$-(1/q)J_{er}(a) = (1/q)J_{hr}(a) = sn_1(a), \quad (18)$$

where $J_{er}(a)$ and $J_{hr}(a)$ are respectively the radial components of the electron and the hole current at the surface, s is the surface recombination velocity and $n_1(a)$ is the excess carrier density at the surface. From (1) and (2), for E_0 and B_0 in the z direction and for $\mu_{e,h}{}^2B_0{}^2\ll 1$, we have to first order in the perturbation quantities

$$-\frac{1}{q}J_{er} = n_0\mu_e\left(\frac{\partial\psi_1}{\partial r} - \frac{\mu_e B_0}{r}\frac{\partial\psi_1}{\partial\phi}\right) - D_e\left(\frac{\partial n_1}{\partial r} - \frac{\mu_e B_0}{r}\frac{\partial n_1}{\partial\phi}\right) \quad (19)$$

and

$$-\frac{1}{q}J_{hr} = -p_0\mu_h\left(\frac{\partial\psi_1}{\partial r} + \frac{\mu_h B_0}{r}\frac{\partial\psi_1}{\partial\phi}\right) - D_h\left(\frac{\partial n_1}{\partial r} + \frac{\mu_h B_0}{r}\frac{\partial n_1}{\partial\phi}\right). \quad (20)$$

Substituting these expressions into (18) and introducing the solutions (11), (12), and (17) into the resulting equations we obtain two linear, homogeneous equations for the coefficients c_1 and c_2. The dispersion relation relating the frequency ω and propagation constant k is obtained by setting the determinant of these equations equal to zero. The result, after inserting the value of A from (15) and performing some algebraic manipulation,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is the following equation:

$$(m\mu_M^2 E_0 B_0/D_a)[k/(\beta^2-k^2)]$$
$$\times[\beta I_m'(\beta a)I_m(ka)-kI_m(\beta a)I_m'(ka)]$$
$$+a\beta kI_m'(\beta a)I_m'(ka)-im\mu_H B_0\beta I_m'(\beta a)\,I_m(ka)$$
$$+[(sa/D_a)-im\mu_a B_0]kI_m(\beta a)I_m'(ka)$$
$$-im\mu_H B_0(s/D_a)I_m(\beta a)I_m(ka)=0, \quad (21)$$

where μ_a, the ambipolar mobility, is as defined in (10),

$$\mu_H=R_H\sigma_0=(p_0\mu_h^2-n_0\mu_e^2)/(n_0\mu_e+p_0\mu_h) \quad (22)$$

is the mixed conduction Hall mobility, and

$$\mu_M^2=n_0p_0\mu_e\mu_h(\mu_e+\mu_h)^2/(n_0\mu_e+p_0\mu_h)^2 \quad (23)$$

is a quantity with the dimensions of mobility squared which appears also in the theory of transverse magnetoresistance. In obtaining (21) use has been made of the definition (10) of D_a and, consistent with the earlier approximation $\mu_{e,h}^2B_0^2\ll1$, a term in B_0^2 has been dropped.

This dispersion relation (21) is rather complex and extremely difficult to treat as it stands. Besides being transcendental, it contains both ω and k in the arguments of the Bessel functions (through the quantity β). Fortunately, for much of the region of interest it is a reasonable approximation to expand the Bessel functions in Taylor series to obtain an algebraic equation for ω and k. In the section on experimental results some numerical solutions of (21) will be presented which indicate the range of validity of these expansions. Further, since we are dealing with the surface-density mode, it is expected that surface recombination will have to be small in order to obtain growth. Therefore, we make the approximation that $s\ll D_a/a$ and neglect the terms involving s. In addition, it is assumed that bulk recombination is also sufficiently small so that $\tau_b\ll a^2/D_a$, which makes it unnecessary to reinsert the bulk recombination term. In Appendix B the modifications due to recombination are taken into account and it is found that for the experimental situation of interest the correction is negligible.

For simplicity we consider here only the lowest order $|m|=1$ mode as this mode has the lowest threshold for growth. In Appendix C the general m mode is considered. Expanding the Bessel functions, we keep only the first two terms of the Taylor series, which will be valid as long as β^2a^2 and $k^2a^2\ll20$. Substituting the expansions in (21), neglecting terms in β^2k^2, and eliminating β^2 by means of (13) we obtain for the simplified dispersion relation:

$$k^2a^2[3+2i(\mu_e-\mu_h)B_0]+i(a^2/2D_a)$$
$$\times(\omega-\mu_aE_0k)[3-i(3\mu_H+\mu_a)B_0]$$
$$+(\mu_M^2E_0B_0/D_a)ka^2+4[1+i(\mu_e-\mu_h)B_0]. \quad (24)$$

Threshold Conditions

As discussed in Sec. II, it is expected that for a given electric field there should be a critical magnetic field and a critical frequency at which growth is first observed. At this point, which we shall call the threshold, the frequency ω and propagation constant k are both real, corresponding to neither growth nor attenuation. Therefore, at threshold the real and imaginary parts of (24) may be separated and ω eliminated between the resulting equations. Neglecting terms of order $\mu_{e,h}^2B_0^2$ compared with unity the result is

$$3k^2a^2+(\mu_M^2E_0B_0/D_a)ka^2+4=0. \quad (25)$$

It is evident that only the product E_0B_0 enters (25). Since we seek the minimum value of the applied fields for which growth is first observed we minimize the E_0B_0 product with respect to k. In this manner we obtain from (25) the threshold value of k, denoted by k_c:

$$k_c^2=4/3a^2, \quad (26)$$

and also the threshold value of E_0B_0. For ease in comparing theory with experiment, the latter is written as an expression for the threshold magnetic field B_{0c}, which is then a function of E_0:

$$B_{0c}=-\frac{6k_cD_a}{m\mu_M^2}\frac{1}{E_0}, \quad (27)$$

with $|m|=1$.

The critical frequency f_c is obtained from (24). Since at threshold $E_0\sim B_{0c}^{-1}$ from (27), we must consider two limiting cases in order to remain consistent with the assumption $\mu_{e,h}^2B_0^2\ll1$. For extrinsic material, such that $|(n_0-p_0)/(n_0+p_0)|\gg\mu_{e,h}^2B_0^2$, all terms except the one involving μ_a are neglected, so that

$$f_c=(1/2\pi)\mu_aE_0k_c. \quad (28)$$

For intrinsic or nearly intrinsic material, where $|(n_0-p_0)/(n_0+p_0)|\ll\mu_{e,h}^2B_0^2$, only the terms in $\mu_{e,h}B_0$ are retained, with the result

$$f_c=-(20mD_a/9\pi a^2)(\mu_e-\mu_h)B_{0c}. \quad (29)$$

Equation (28) corresponds to the translation of the helix at the ambipolar drift velocity (μ_aE_0), while (29) corresponds to the rotation of the helix due to diffusion and conduction across the magnetic field. In the intermediate case $|(n_0-p_0)/(n_0+p_0)|\approx\mu_{e,h}^2B_0^2$, the total motion is composed of both translation and rotation; however, a correct expression for f_c cannot be obtained in the approximation $\mu_{e,h}^2B_{0c}^2\ll1$.

From (26) it is evident that just the magnitude of k_c is determined. This is simply a consequence of the fact that for a wave of the form $\exp(i\omega t-ikz-im\phi)$ only the relative signs of k, m, and ω are significant. We will choose the signs of k_c and m to make $f_c>0$ in (28) or (29) and to satisfy (27). The resulting signs of ω, k, and m for various directions of E_0 and B_0 and for different types of material are listed in Table I. We have assumed $\mu_e>\mu_h$. The sense and motion of the helix may be deduced immediately from this table.

In agreement with the predictions of the physical

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Relative signs of ω, k, and m for helical wave of proper sense to permit growth.

Directions of E_0 and B_0	n-type material	p-type material	Intrinsic material
$E_0>0$ $B_0>0$	$\omega>0$ $k>0$ $m<0$	$\omega>0$ $k<0$ $m>0$	$\omega>0$ $k>0$ $m<0$
$E_0>0$ $B_0<0$	$\omega>0$ $k>0$ $m>0$	$\omega>0$ $k<0$ $m<0$	$\omega>0$ $k>0$ $m>0$
$E_0<0$ $B_0<0$	$\omega>0$ $k<0$ $m>0$	$\omega>0$ $k>0$ $m<0$	$\omega>0$ $k<0$ $m>0$
$E_0<0$ $B_0>0$	$\omega>0$ $k<0$ $m<0$	$\omega>0$ $k>0$ $m>0$	$\omega>0$ $k<0$ $m<0$

model, the helix that can exhibit growth is right- (left-) handed for B_0 parallel (antiparallel) to E_0, irrespective of the type of material. In each case the screw of opposite sense has no threshold and is always attenuated, as will be seen in a subsequent calculation of the growth constant. In any cross-sectional plane, the direction in which the helix appears to revolve about its axis is independent of the direction of E_0 and reverses sign with a reversal of B_0. Further, for a given E_0 and B_0 this direction of apparent revolution is the same for n-type and intrinsic material and is in the opposite direction for p-type material.

Growth and Phase Constants

Having established that a threshold for growing helical waves does indeed exist and having calculated the critical fields, wave number, and frequency, we now examine in detail the functional behavior of the propagation constant of the waves both above and below threshold. In particular we will consider sufficiently extrinsic material that the waves exhibit only stable spatial growth. The question of instability will be taken up in detail later and it will be seen just how extrinsic the material must be to assure stable growth.

We seek solutions of the dispersion relation in the form of traveling waves, excited at a specific frequency and varying spatially. Therefore, ω is considered to be a real variable and k is allowed to become complex, i.e., $k=k_r+ik_i$, where k_r is the phase constant and k_i the growth constant. The wave will then travel with phase velocity $v_{ph}=\omega/k_r$ and will vary in amplitude as $\exp(k_i z)$. Equation (24) is now solved directly for k, and for material sufficiently extrinsic that

$$\mu_a E_0 \gg D_a/a\,, \tag{30a}$$

$$|(n_0-p_0)/(n_0+p_0)| \gg \mu_{e,h}{}^2 B_0^2\,, \tag{30b}$$

and

$$(\mu_a E_0)^2 \gg \omega D_a\,, \tag{30c}$$

the square root in the quadratic formula can be expanded in a Taylor series. In performing the expansion we retain real terms to first order and imaginary terms to second order. Further, we choose the root corresponding to the negative square root; the other choice leads to $k_i \approx \mu_a E_0/2D_a$ which by (30a) is much greater than unity and is thus not self-consistent with the earlier Bessel function expansion. Actually the full transcendental dispersion relation leads to an infinite number of other roots, which presumably correspond to strongly damped modes. For the correct root we have

$$k \approx (\omega/\mu_a E_0) - i(8D_a/3a^2\mu_a E_0)[1+\tfrac{3}{4}(\omega a/\mu_a E_0)^2 + (\omega a^2 \mu_M{}^2 B_0/4\mu_a D_a)]. \tag{31}$$

A considerable simplification of this result may be achieved through the use of the previously derived expressions (27) and (28) for the threshold frequency and magnetic field. The propagation constant $k=k_r+ik_i$ is then

$$k \approx \frac{\omega}{\mu_a E_0} + i\frac{8D_a}{3a^2\mu_a}\frac{1}{E_0}\left[2\frac{fB_0}{f_c B_{0c}} - \left(\frac{f}{f_c}\right)^2 - 1\right]. \tag{32}$$

Equation (32) corresponds to the helical wave of proper sense to permit growth, as given by Table I. The propagation constant for the wave of opposite sense is obtained by simply reversing the sign of the term $(2fB_0/f_c B_{0c})$.

Since $k_r \gg k_i$, we have from (32)

$$v_{\text{phase}} = v_{\text{group}} = \mu_a E_0\,, \tag{33}$$

confirming our earlier assertions that for extrinsic material the wave translates with the ambipolar drift velocity.

We see further that the wave will grow spatially ($k_i>0$ for waves traveling in the direction of increasing z and $k_i<0$ for those propagating in the opposite direction) when

$$2B_0/B_{0c} > (f/f_c)+(f_c/f) \tag{34}$$

and the wave of opposite sense is always attenuated.

Actually the existence of a value of k_i of the correct sign for growth is not sufficient to prove that the solution corresponds to a growing wave; it might also be an evanescent wave, such as in a waveguide beyond cutoff. However, the experimental results confirm the fact that (32) is a true growing wave solution, and we need not investigate this point further.

Instability

In the earlier discussion of the physical model for the helical wave it was reasoned that for sufficiently small ambipolar drift velocity a disturbance could grow temporally rather than spatially, resulting in instability as opposed to stable growth. We shall now derive this result mathematically by developing a quantitative stability criterion for the surface density mode of operation.

Stability criteria and the important distinction between convective growth and absolute instability in systems with growing wave behavior have been discussed by several authors.[14–19] The problem is essentially one of determining the response of the system to a spatially and temporally bounded, but otherwise arbitrary, excitation. If the response (which may be any physical quantity of interest) is bounded as $t \to \infty$ then the system is stable, and it is absolutely unstable otherwise. Assuming that only one dimension need be considered and

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that end boundaries are sufficiently far removed so that reflected waves are of no significance, it can be shown that for the purposes of determining the stability it is sufficient to investigate the response $\phi(z,t)$ to a δ-function excitation in space and time:

$$\phi(z,t)=\int_{C_L} e^{i\omega t}d\omega \int_{C_F} \frac{e^{-ikz}}{D(k,\omega)}dk, \tag{35}$$

where $D(k,\omega)=0$ is the dispersion relation for the system under consideration. The Laplace contour C_L runs in the lower half of the ω plane, below all singularities of the integrand, i.e., $-\infty-i\nu<\omega<\infty-i\nu$, where ν is a large positive number. For convenience we have set $s=i\omega$ in the conventional Laplace formulation. The Fourier contour C_F is as usual along the real k axis. For positive t the ω-integration contour is closed by a semicircle at infinity in the upper half-plane, and the k contour is closed in the upper or lower half-plane when the region of interest is for $z<0$ or $z>0$, respectively.

In general the integrals in (35) will be difficult and several authors[14–19] have proposed techniques for obtaining the stability information without explicitly evaluating these integrals. However, for the simplified helical-wave dispersion relation (24) for the surface-density mode, the integrations in (35) can be performed directly to investigate the behavior of $\phi(z,t)$ for large t. From the physical model presented in Sec. II it is expected that instability will be possible only for nearly equal densities of positive and negative carriers, so for simplicity we set $n_0=p_0$ in most terms of (24). However, we do not set $\mu_a=0$, and thereby retain the very important terms involving the quantity (n_0-p_0). This turns out to be equivalent to neglecting terms of the order of $(n_0-p_0)/(n_0+p_0)$ with respect to unity, but not with respect to quantities of order $\mu_{e,h}B_0$, which are much less than unity. This level of approximation will be maintained throughout the calculation. In several places we shall neglect to set $n_0=p_0$ purely to preserve the form of particular terms and to aid in their identification later; however, it is to be understood that the following calculations are correct only for $n_0\approx p_0$.

Equation (24), after rationalization of the leading term, may then be written in the form (assuming $\mu_{e,h}^2B_0^2\ll 1$ as usual):

$$D(k,\omega)=(k-k_0)^2-\alpha(\omega-\omega_0)=0, \tag{36}$$

where

$$k_0\equiv -m\mu_e\mu_hE_0B_0/6D_a+i[\mu_a/4D_a+\mu_e\mu_h(\mu_e-\mu_h)B_0^2/9D_a]E_0, \tag{37a}$$

$$\omega_0\equiv -(m\mu_e\mu_hE_0^2B_0/6)[\mu_a/D_a-5\mu_e\mu_h(\mu_e-\mu_h)B_0^2/3D_a]+i[8D_a/3a^2+(\mu_aE_0)^2/8D_a-\mu_M^4E_0^2B_0^2/18D_a]. \tag{37b}$$

$$\alpha\equiv[m(\mu_e-\mu_h)B_0-3i]/6D_a. \tag{37c}$$

Focusing our attention on the Fourier integral of (35) in the k plane, we see that only contributions to the integral will be at values of $k=k(\omega)$ such that $D(k,\omega)=0$. As the notation implies, these values of k depend parametrically on ω, where ω assumes all values on the Laplace contour C_L. Referring to (36) it is clear that there will be two lines of zeros of $D(k,\omega)$ in the k plane, corresponding to the two solutions

$$k_\pm(\omega)=k_0\pm[\alpha(\omega-\omega_0)]^{1/2}, \tag{38}$$

where

$$\omega=\omega(C_L). \tag{39}$$

Since C_L is below all singularities of the integrand in the ω plane, we may immediately conclude that the two lines of zeros do not cross the real k axis and that all of the zeros are simple. Further, using (37c) and (38) in (36) and recalling that ν may be an arbitrary large positive number, it is clear that $k_+(\omega)$ corresponds to the line of zeros in the upper half k plane, and $k_-(\omega)$ to that in the lower half-plane.

Since the zeros of $D(k,\omega)$ are simple, the poles of the integrand at $k=k_\pm(\omega)$ are also simple and the integral over k can be evaluated using the theory of residues. For $z>0$ the contour is closed in the lower half-plane enclosing the line of zeros corresponding to $k=k_-(\omega)$. The solution for $z<0$ proceeds analogously and will not be detailed. We may thus write for the response:

$$\phi(z,t)=\exp(-ik_0z)\int_{C_L}\frac{\exp\{i\omega t+i[\alpha(\omega-\omega_0)]^{1/2}\}}{[\alpha(\omega-\omega_0)]^{1/2}}d\omega. \tag{40}$$

The integral in (40) may be evaluated directly from a table of Laplace transforms,[27] with the result

$$\phi(z,t)=\frac{1}{(\alpha t)^{1/2}}\exp\left(i\omega_0t-i\frac{a}{4t}-ik_0z\right). \tag{41}$$

Writing $\omega_0=\omega_{0r}+i\omega_{0i}$, we immediately conclude that the system will be stable or absolutely unstable depending on whether ω_{0i} is positive or negative.

It therefore follows from (37b) that a long semiconductor bar with a low-recombination surface and with applied parallel electric and magnetic fields E_0 and B_0 will be stable ($\omega_{0i}>0$) when

$$(\mu_aE_0)^2+(8D_a/a\sqrt{3})^2\geq(\tfrac{2}{3}\mu_M^2E_0B_0)^2 \tag{42}$$

and will exhibit absolute instability when the inequality is reversed.

We may now make a plausible identification of the terms of (42). The quantity μ_aE_0 is simply the ambipolar drift velocity. The term $(8D_a/a\sqrt{3})$ is in the form of a diffusion velocity and is a measure of the effective velocity at which carriers diffuse out of the helix. To aid in the identification of the last term we note that the earlier results on the threshold for growth may be put in a form similar to (42). From (26) and (27) we may write that growth will obtain when

$$(\tfrac{2}{3}\mu_M^2E_0B_0)^2>(8D_a/a\sqrt{3})^2. \tag{43}$$

[27] See, for example, R. V. Churchill, *Operational Mathematics* (McGraw-Hill Book Company, Inc., New York, 1958), p. 328.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The right-hand term is again the square of the effective diffusion velocity. Since it is diffusion which must be overcome to obtain growth, we may conclude that the left-hand term is the square of the effective velocity at which carriers are driven into the helix by the combined action of E_0 and B_0, corresponding to the growth mechanism.

With these identifications in mind we may write the three quantities in (42) as v_{drift}^2, v_{diff}^2, and v_{growth}^2, respectively. From (42) and (43) it may then be stated that the helical wave will exhibit stable traveling-wave amplification when

$$v_{\text{drift}}^2+v_{\text{diff}}^2>v_{\text{growth}}^2>v_{\text{diff}}^2, \tag{44}$$

and will be unstable when the first inequality is reversed.

Although the above discussion is based on a plausible and not rigorously justified identification of terms in (42) and (43), we see that (44) leads us to essentially the same conclusions reached by physical reasoning. That is, when the flow of particles due to the growth mechanism is sufficiently strong to overcome diffusion, a disturbance will grow but will be swept away from its point of origin and spatial growth will result. Further, when the flow is sufficient to overcome both diffusion and drift, the disturbance will build up temporally at its point of origin, resulting in instability. It is not clear, however, why the above expressions involve the squares of the various velocities. Although these velocities are actually complex vectors, and some components will therefore be in either or both time and space quadrature, no explanation of the form of (44) has yet been obtained on this basis.

Several further interesting conclusions may be drawn from Eqs. (42)–(44). For $n_0\neq p_0$ the threshold for growth is lower than that for instability, and consequently as B_0 is increased for fixed E_0 the system passes successively from attenuated waves to stable growing waves to instability. For n_0 exactly equal to p_0, however, there is no region of stable gain and above threshold the system is immediately unstable. The instability threshold value of E_0 decreases monotonically with increasing B_0, but there is a minimum value of B_0 given by

$$B_{0,\min}=3\mu_a/2\mu_M^2 \tag{45}$$

below which there can be no instability regardless of the value of E_0. The instability threshold value of B_0 also decreases monotonically with E_0 and although (42) predicts no corresponding minimum E_0 for instability, there presumably is such a minimum due to saturation of the growth term when $\mu_{e,h}B_0\gg1$. The terms which lead to this saturation have, of course, been neglected throughout the calculations.

Finally, we may calculate the frequency and wave number at which instability is first predicted, i.e., when $\omega_{0i}=0$ and (42) becomes an equality. These quantities are just ω_{0r} and k_{0r}, respectively, and are given by (37a) and (37b), where E_0 and B_0 are the instability threshold values and are related by (42) with the inequality sign replaced by an equal sign.

IV. EXPERIMENTAL PROCEDURE

In order to provide quantitative verification of the theory presented in the preceding section, a long, uniform cylindrical semiconductor sample was required with accurately known carrier densities and mobilities and with bulk and surface recombination sufficiently low that $\tau_b\gg a^2/D_a$ and $s\ll D_a/a$. Further, it was necessary that the longitudinal electric field, pulsed to avoid heating, be applied by means of truly Ohmic contacts in order not to alter the uniform equilibrium carrier distribution. Finally, a means of exciting and detecting the helical wave without altering the number of minority carriers was required.

The samples were fabricated from high-quality, uncompensated n-type germanium with a room-temperature resistivity of about 30 ohm-cm and a bulk lifetime of 1400 μsec. Accurate values for the carrier densities at each temperature of interest (290–360°K) were computed from the value of $n-p=N_D-N_A=3.16\times10^{13}$, obtained by Hall measurements at 77°K, plus the reported value[28] of the np product. In order to ensure that the net donor density was given correctly by the usual expression $N_D-N_A=-(R_Hq)^{-1}$, where R_H is the Hall coefficient, the measurements were made with a magnetic field of 20 kG oriented in a ⟨001⟩ crystallographic direction.[29] Further, the Hall sample had a length-to-width-ratio of 6 to prevent shorting of the Hall field by the end contacts.[30] The uncertainty in the resulting value of N_D-N_A is estimated to be $\pm3\%$.

Carrier-drift mobilities were obtained by scaling the reported[31] lattice mobilities at 300°K according to the experimentally obtained[32] power law appropriate to each carrier for lattice scattering; i.e., $\mu_e=(3900)(T/300)^{-1.66}$ and $\mu_h=(1900)(T/300)^{-2.33}$. This computation was felt to be fully valid in view of the fact that in high-quality uncompensated germanium, lattice scattering is completely dominant in this temperature range.[31,32] Further, the measured resistivities at each temperature (using voltage probes and electric fields of about 0.05 V/cm to avoid complications due to contacts) agreed well within experimental error with the values calculated from the above densities and mobilities.

Samples of both square and circular cross section were used. The samples were first lapped and then etched to final size to remove all damaged surface layers. Dimensions after etching were approximately 25 mm long$\times$1 mm diam for the cylindrical samples and 25$\times$1$\times$1 mm for those of square cross section. To comply with the requirement of the theory, the samples should be much longer than the attenuation length of

[28] F. J. Morin and J. P. Maita, Phys. Rev. 94, 1525 (1954).
[29] W. M. Bullis and W. E. Krag, Phys. Rev. 101, 580 (1955).
[30] C. Dunlap, *An Introduction to Semiconductors* (John Wiley & Sons, Inc., New York, 1957), p. 186.
[31] M. B. Prince, Phys. Rev. 92, 681 (1953).
[32] F. J. Morin, Phys. Rev. 92, 62 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

any reverse wave so that the presence of the ends cannot be communicated to a significant portion of the bar. Since it is not known theoretically if a reverse wave does exist and if so what its properties are, the sample length was arbitrarily made approximately ten times the threshold wavelength [see Eq. (26)]. Experimentally, no significant effects due to end boundaries were observed.

The etching procedure was quite complex and has been described in detail elsewhere.[33] Using this procedure, which was concluded by protecting the surfaces with silicone oil, it was possible to achieve and maintain surface recombination velocities which were regularly less than 50 cm/sec and often as low as 20 cm/sec. These values of s were computed[34] from the lifetime of injected carriers, measured by decay of photoconductivity for each sample, and the similarly measured bulk lifetime of the original crystal.

The relatively high-resistivity material and strong electric fields (20–60 V/cm) precluded the use of the usual methods for making Ohmic contacts. Metallic soldered, plated, or alloyed contacts always gave p^+n or n^+n junctions which interfered with the flow of carriers. Depending on the direction of current flow a p^+n contact injected or extracted holes while an n^+n contact excluded or accumulated them. Sandblasting the surfaces near the contact to recombine an excess or make up a deficit of minority carriers failed because the electric field swept the carriers through any practical recombination-generation region much faster than they could diffuse to or from the surface.

The problem was solved by use of the arrangement shown in Fig. 2. Only one end of the bar is labeled but the other end is identical. The p^+ contact carries only the hole current and the parallel n^+ contacts carry only the electron current. The series resistors are much larger than the contact resistances. Thus, by adjusting the variable resistor in series with the p^+ contact, the relative electron and hole currents can be set to precisely the values required to maintain the thermal-equilibrium carrier distribution. The large resistors make the setting independent of the electric field, but of course the variable resistor has to be readjusted for each temperature of operation. With respect to fabrication, excellent contacts were formed by alloying In+1/2% Ga and Sn+5% Sb spheres to the sample followed by electrolytic etching.[33]

Excitation and detection of the waves were accomplished by means of pairs of small n^+ probes attached to the surface of the sample (see Fig. 2). These n^+n junctions conduct majority carriers (electrons) freely but are impervious to the flow of minority carriers (holes).[35]

Thus when a pair of probes (such as aa′) is driven

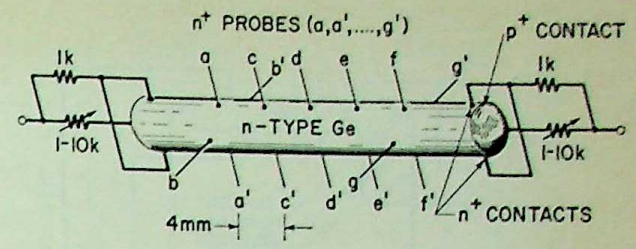


FIG. 2. Sketch of cylindrical sample, showing Ohmic end contacts and n^+ probes.

with an ac signal source, the electric field produced causes the holes to move back and forth in synchronism, perturbing the carrier distribution and initiating the wave. Actually, the linearly polarized excitation just described excites both the $m=\pm 1$ modes, but as discussed in the last section, one mode will always decay while the other can grow. By using two sets of probes (aa′ and bb′, for example) which are 90° apart around the circumference and driving them with two signals 90° out of time phase, one or the other of the circularly polarized modes may be excited preferentially.

The use of the n^+ probes in detecting the helical wave makes use of the fact that the electron quasi-Fermi level (QFL) is constant through the n^+n junctions. The open-circuit voltage between two probes then gives the difference between the electron QFL's just beneath the surface at the two points. However, for small signals the perturbed electron QFL is linearly related[36] to the perturbations in the density and potential, i.e.,

$$\phi_{e1}\approx\psi_1-(kT/q)(n_1/n_0).$$

Therefore, all information pertaining to the amplitude and phase characteristics of the waves may be obtained by measuring ϕ_{e1} with n^+ probes. Of course, it is also necessary that the contacts be impervious to the flow of holes in order to prevent their generation or recombination in the junctions.

These probes were fabricated by bonding 0.002-in. Sb-doped Au wires to the sample with pulses of current.[37] To assure that the two probes of a set (c and c′, for example) were aligned on the same cross-sectional plane of the sample, an electrical null technique was used to locate them.[33] All probes were checked for proper operation by using them to measure known changes in the electron QFL produced by illuminating the sample with penetrating light.[33] This test was also correlated with the current-voltage characteristics of the probes.

A diagram of the arrangement for the above-described excitation and detection of the waves is shown in Fig. 3. The phase shifter and probe driver provided signals 90° out of phase from low-impedance sources. The input impedance of the detecting and measuring circuits was

[33] C. E. Hurwitz, MIT Lincoln Laboratory Technical Report No. 320, 1963 (unpublished).
[34] W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., New York, 1950), pp. 318–325.
[35] A. K. Jonscher, *Principles of Semiconductor Device Operation* (G. Bell and Sons, London, 1960), pp. 142–144.
[36] W. Shockley, *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, Inc., New York, 1950), p. 308.
[37] M. C. Waltz, in *Transistor Technology*, edited by H. E. Bridges, J. H. Scaff, and J. N. Shive (D. Van Nostrand Company, Inc., Princeton, 1958), Vol. I, pp. 375–380.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

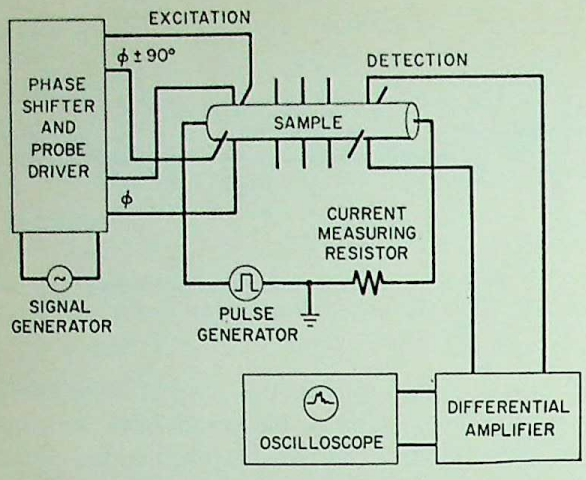


FIG. 3. Experimental arrangement for exciting and detecting helical waves. The differential amplifier may be connected to any desired set of probes.

sufficiently high to provide "open circuit" conditions up to frequencies in excess of 500 kc.

The completed sample, mounted in a jig, was attached to a holder which permitted rotation of the sample in a vertical plane for accurate alignment parallel to the magnetic field. This arrangement was mounted between the poles of an electromagnet which was rotatable in the horizontal plane to complete the alignment of sample and field. The sample and lower part of the holder were immersed in a bath of silicone oil which was maintained at a constant, controllable temperature. A thermocouple mounted on the jig next to the sample read the actual temperature. In this manner the temperature could be set and held to within $\pm 0.1°C$ of the desired value.

V. EXPERIMENTAL RESULTS AND COMPARISON WITH THEORY

The basic experiment consisted of applying parallel dc-magnetic and pulsed-electric fields to the sample, exciting the wave with a signal source at an appropriate set of probes, and then observing the excitation in magnitude and phase as it traveled down the bar.

Many samples of both square and circular cross section were fabricated and examined experimentally. All exhibited spatially growing waves whose properties agreed qualitatively with the predictions of the physical model and the mathematical theory. However, due to the involved fabrication procedure and the finite-failure rate at each step (a full set of properly operating n^+ probes was particularly difficult to achieve), only two samples, one square and one circular, were judged to fully meet the requirements discussed in the previous section and therefore to be suitable for accurate quantitative measurements. As will become evident in the following discussion, the agreement between experiment and theory is sufficiently good that although results for only two samples are reported, the theory can be considered completely verified in the region of stable growth.

The final dimensions of these two samples were 1.16 mm diam by 25.4 mm long for the cylindrical sample and $1.02 \times 1.02 \times 25.4$ mm for the one of square cross section.

The temperature of operation proved to be a particularly valuable parameter to vary as it provided a test of the functional dependence of the wave properties on the equilibrium carrier densities without the need for many samples of varying resistivities. In fact, by raising the temperature it was possible to pass from the extrinsic region in which stable spatial growth is observed to nearly intrinsic material where some of the considerations of instability and the oscillistor are of interest.

The extremely low yield of good samples precluded the use of sample geometry as another variable. The dependence on sample radius seems well enough confirmed by the general quantitative agreement of all measurements, but the question of how long is a "long" sample and the length dependence of the threshold in samples comparable in size to a wavelength would still be of interest. Except where otherwise stated, the data to be presented refer to the sample of circular cross section, for in this case there is no ambiguity as to the value of radius to be used in the theoretical expressions.

Helical Waveshape

The helical nature of the wave was confirmed by exciting the sample at probes aa′ and observing the relative phase of the resulting signals at probes ff′ and gg′. (See Fig. 2. The electric field is positive from left to right.) The resulting waveforms are shown in Fig. 4, and we see that for B_0 parallel to E_0 the signal at gg′ led that at ff′ by 90° and lagged by 90° when B_0 was reversed. The traces shown are for threshold, but the same phase relationship was found for all fields and frequencies. Adding to this result the fact that the phase and group velocities are in the direction of the electric field (see next section) it is clear that we do indeed have a helical wave, traveling in the direction of ambipolar drift, which as predicted is right- (left-) handed for B_0 parallel (antiparallel) to E_0.

Further, when a circularly polarized excitation was provided by driving probes aa′ and bb′ with signals 90° out of time phase, only the polarization direction cor-

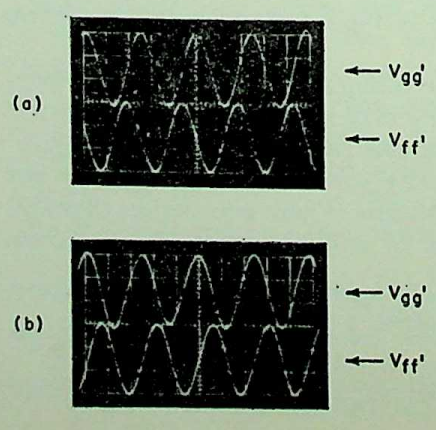


FIG. 4. Oscilloscope traces showing 90° phase difference between signals at probes gg′ and ff′. Set (a) corresponds to B_0 parallel to E_0 and (b) to B_0 antiparallel to E_0. Vertical scale: 5 mV/div. Horizontal scale: 10 μsec/div.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

responding to the predicted screw sense produced a growing wave. The opposite excitation produced only a strongly attenuated wave. That is, for excitation in the sense $a \rightarrow b' \rightarrow a' \rightarrow b$ the wave grew (for fields above threshold) for B_0 parallel to E_0 and always decayed when B_0 was reversed. For excitation in the sense $a \rightarrow b \rightarrow a' \rightarrow b'$ the converse was true. These results are in full agreement with the predictions of the physical model, Eq. (32) and Table I.

Phase Constant—Phase and Group Velocity

For a wave of the form $\exp[i\omega t - i(k_r + ik_i)z \pm i\phi]$ the phase difference $\Delta\Phi$ between two points differing in z by Δz at the same instant of time and same value of ϕ is simply

$$\Delta\Phi = -k_r \Delta z.$$

Thus the measurement of the relative phase of the signal at two sets of probes, such as dd′ and ff′, gives directly the phase constant k_r. In Fig. 5 we plot $-\Delta\Phi$, the amount by which the phase at ff′ lags that at dd′, and the resulting value of k_r, as a function of f/E_0 for several temperatures and electric fields. The probe spacing Δz in this case was 0.8 cm. Since the probes were several wavelengths apart there was always an uncertainty of $2n\pi$ in the absolute phase difference as related to that measured. However, for the extrinsic case we assume that we have a simple wave so that the extrapolated value of $\Delta\Phi$ is zero for $\omega = 0$. Therefore, to each set of experimental points (at a given temperature) we have added an appropriate multiple of 2π. For each set of points the required translation was always $2n\pi$ and n was just that integer required to make $k_r = 2/a\sqrt{3}$ at threshold, in agreement with (26).

The measured values of $\Delta\Phi$ and therefore k_r were, within experimental error, independent of magnetic field over the range 0–11 kG for all temperatures except 67°C and above, but in the latter range exhibited a very slight dependence at the lowest frequency and highest

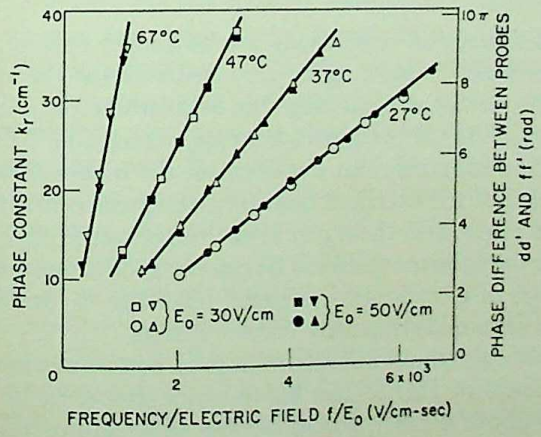


FIG. 5. Phase constant k_r (left-hand scale), computed from measured phase difference between probes dd′ and ff′ (right-hand scale), versus f/E_0, for several temperatures and electric fields. Points are experimental; curves are theoretical, calculated from Eq. (32).

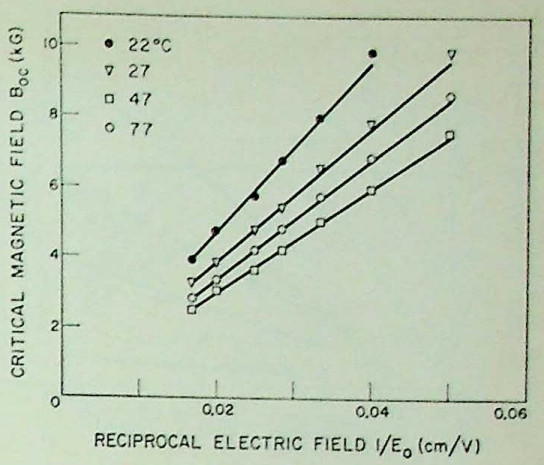


FIG. 6. Critical magnetic field B_{0c} versus $1/E_0$ for several temperatures. Points are experimental; curves are theoretical, calculated from Eq. (27).

magnetic fields. The theoretical curves were calculated from (32) and agreement is excellent. The slopes of the straight lines are just $\mu_a/2\pi$ and therefore, as predicted, $v_{\text{phase}} = v_{\text{group}} = \mu_a E_0$, the ambipolar drift velocity. Values of k_r obtained by numerical computation from the full dispersion relation (21) were insignificantly different from those given by (32).

Threshold Conditions

The threshold, by definition, is the point at which the z variation of the wave first changes from attenuation to growth. At this point $k_i = 0$ and the signal at each set of detecting probes must have the same magnitude. For a fixed electric field, the threshold frequency f_c and magnetic field B_{0c} were obtained by exciting the wave at probes aa′ and bb′ and then slowly raising the magnetic field while sweeping the frequency back and forth. The field and frequency at which the signals at probes dd′, and ee′, and ff′ first became equal were then the threshold values. For properly operating probes these signals became equal simultaneously. The signal at probes cc′ was usually slightly larger, presumably due to higher modes which had not decayed sufficiently that near to the excitation.

In Figs. 6 and 7 we have plotted the measured values of f_c and B_{0c} as a function of the applied electric field E_0 for various temperatures. The theoretical curves were calculated from (27) and (28). Agreement is again very good. The small deviations from theory of the values of f_c at the higher temperatures, lowest electric fields, and highest magnetic fields are presumably due to the beginning influence of the rotational contribution to the frequency. At high magnetic fields, B_{0c} also begins to deviate from theory due to breakdown of the assumption $\mu_{e,h}^2 B_0^2 \ll 1$. Note, however, that at moderate fields the theoretical and experimental values of B_{0c} agree very well at all temperatures. This is to be expected since no assumptions concerning the relative carrier densities were made in the theoretical calculation of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

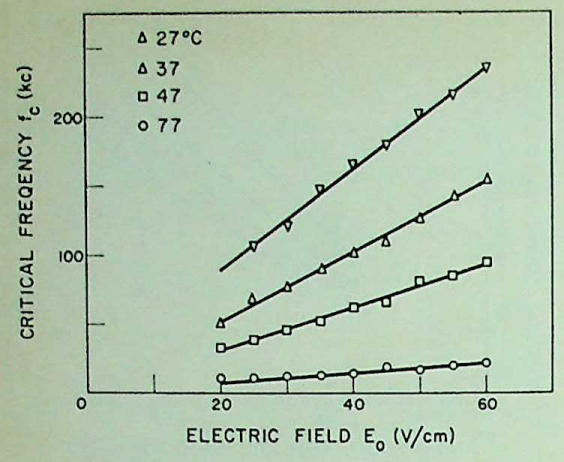


FIG. 7. Critical frequency f_c versus E_0 for several temperatures. Points are experimental; curves are theoretical, calculated from Eq. (28).

B_{0c}. As the temperature increases and $(n_0-p_0)\rightarrow 0$, f_c decreases rapidly, but B_{0c} quickly approaches a limiting value which is independent of the magnitude of n_0 and p_0, as may be seen from Eq. (27). However, when the mobilities are decreasing functions of temperature, as is the case here, B_{0c} reaches a minimum and then increases slightly with temperature. Thus, as we see in Fig. 6, the values of B_{0c} at 77°C are somewhat higher than at 47°C.[38]

Growth Constant

The exponential variation of the wave amplitude in the z direction was confirmed by measurement of the relative signal magnitudes at probes dd′, ee′, and ff′, with excitation provided at aa′ and/or bb′. More precisely, as long as the excitation was sufficiently small that the signal at any set of probes was less than 500 mV peak-to-peak, the wave varied longitudinally as $\exp(k_i z)$ where k_i is the growth constant. When the signal rose above this level, saturation effects began to set in and the linearized theory no longer applied. More will be said about nonlinear effects in the next section; in the present discussion we limit ourselves to the small-signal linear region.

Figures 8 and 9 show the theoretical and experimental variation of the growth constant k_i as a function of B_0 for various temperatures and values of E_0. In all cases the frequency was fixed at the threshold value f_c pertinent to the specific temperature and electric field. The dashed curves are calculated from (32); the full curves are the results of the exact (in the limit $\mu_{e,h}^2B_0^2\ll 1$) numerical solution of the full dispersion equation, (21). The approximate solution is clearly accurate only in the neighborhood of threshold ($k_i=0$), and deviates both above and below threshold where the

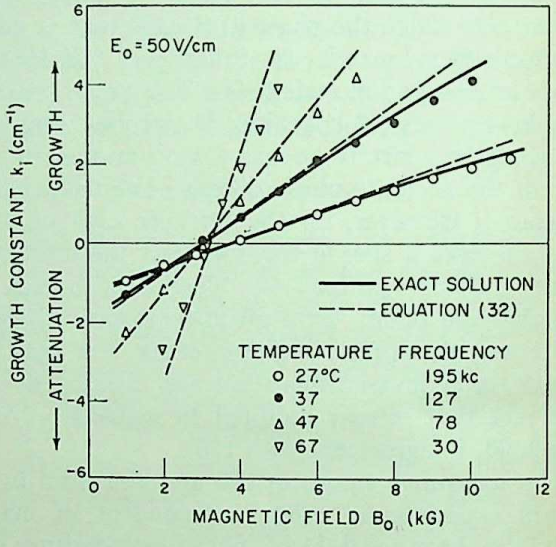


FIG. 9. Growth constant k_i versus B_0 for several temperatures at $E_0=50$ V/cm and $f=f_c$. Points are experimental; solid curves are exact theoretical solutions for $\mu_{e,h}^2B_0^2\ll 1$, and dashed curves are calculated from Eq. (32).

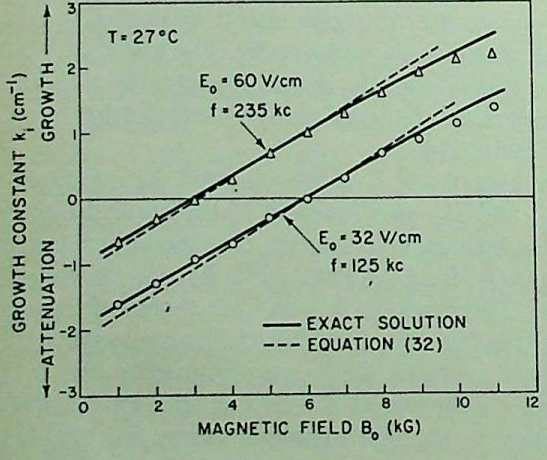


FIG. 8. Growth constant k_i versus B_0 for two values of E_0, at $T=27$°C and $f=f_c$. Points are experimental; solid curves are exact theoretical solutions for $\mu_{e,h}^2B_0^2\ll 1$, and dashed curves are calculated from Eq. (32).

[38] Since all measurements were made with small signals, no hysteresis effects such as described by B. Ancker-Johnson [Appl. Phys. Letters 3, 104 (1963)] were observed.

Bessel function expansions are no longer correct. The experimental results agree very well with the theoretical curves except at high magnetic fields where the assumption $\mu_{e,h}^2B_0^2\ll 1$ becomes a poor approximation. Numerical solutions were not obtained for the higher temperatures, but the excellent quantitative agreement between experiment and theory at lower temperatures indicates that the agreement should be equally good. Intermediate values of electric field, omitted from Fig. 9 for clarity, also gave quantitatively correct results.

Note the rapid increase of growth constant with temperature in Fig. 8. The highest measurable value of k_i was about 4 cm^{-1}, corresponding to a gain of approximately 35 dB/cm. For growth rates above this value the sample began to amplify appreciably its internal noise to values comparable with the maximum signal amplitudes obtainable under linear operation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The frequency dependence of the growth constant is detailed in Figs. 10 and 11, where for $E_0=50$ V/cm the theoretically predicted and measured values of k_i have been plotted as a function of frequency at $T=27°C$ and $T=37°C$ for three values of magnetic field: $B_0<B_{0c}$, $B_0\approx B_{0c}$, and $B_0>B_{0c}$. Again the dashed curves are calculated from (32) and the full curves are the solutions obtained by numerical computation. The inadequacy of the simplified theory, except in the neighborhood of threshold, is strikingly evident. Equally apparent is the extremely good agreement between the experimental points and the exact theory.

In Figs. 12 and 13 similar measurements of k_i as a function of B_0 and f for the sample of square cross section are presented. An effective radius $a=0.053$ cm was chosen in plotting the theoretical curves from (32) in order to give the best fit to the experimental data. This choice is very nearly the radius of a cylinder with the same cross-sectional area as the sample. No numerical solutions were obtained for this case but, in light of the results for the circular sample, the behavior at frequencies and fields above threshold would presumably be in agreement with the exact solution. The gross disagreement at low frequencies, which is more than could be accounted for by the exact theory, is only an apparent one. It is due to the fact that linearly polarized (two-probe) excitation was used rather than the circularly polarized (four-probe) excitation employed for all measurements on the cylindrical sample. Examination of (32) shows that for $f<f_c$ and $B_0<B_{0c}$ the right- and left-handed helices have comparable attenuation rates. Since both are excited approximately equally, the measured value of k_i will be some combination of the two. This identical discrepancy was later noted on the cylindrical sample, but was eliminated by using four-probe excitation. Because the square samples were by then of little interest, no attempt was made to redo the earlier measurements.

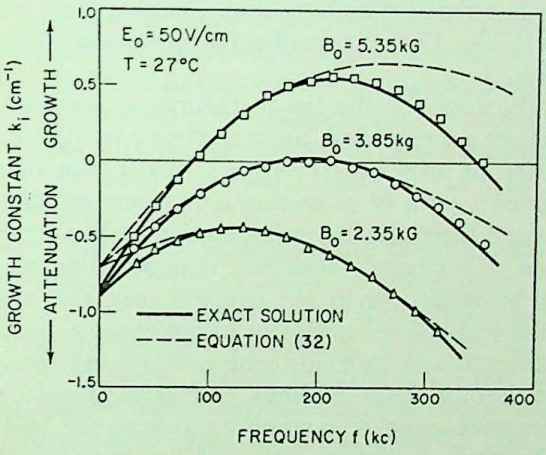


FIG. 10. Growth constant k_i versus f for $B_0<B_{0c}$, $B_0\approx B_{0c}$, and $B_0>B_{0c}$, at $T=27°C$ and $E_0=50$ V/cm. Points are experimental; solid curves are exact theoretical solutions for $\mu_e{}^2B_0{}^2\ll 1$, and dashed curves are calculated from Eq. (32).

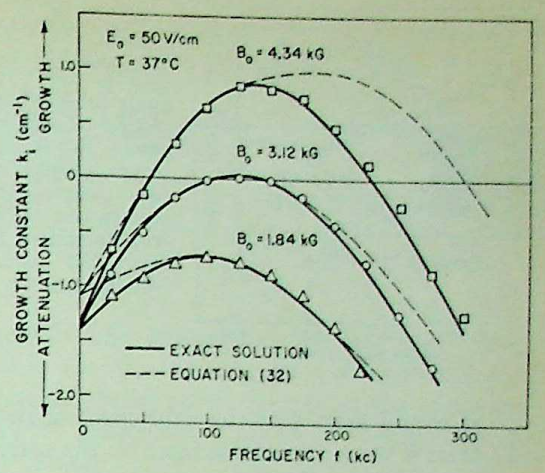


FIG. 11. Growth constant k_i versus f for $B_0<B_{0c}$, $B_0\approx B_{0c}$ and $B_0>B_{0c}$, at $T=37°C$ and $E_0=50$ V/cm. Points are experimental: solid curves are exact theoretical solutions for $\mu_e{}^2B_0{}^2\ll 1$, and dashed curves are calculated from Eq. (32).

An effort to measure directly the functional behavior of k_i for the helix of screw-sense opposite to the one exhibiting growth was made by simply reversing the rotation direction of the four-probe excitation, while maintaining the fields in the same direction. However, due to the fact that four points cannot give a perfect $\exp(i\phi)$ distribution, a sufficient amount of the growing wave was excited to completely obscure the measurements.

Along this same line an attempt was made to excite the next higher $|m|=2$ mode by connecting a and a′ together and b and b′ together, and then driving the resulting two terminals with the signal generator. Such an arrangement should preferentially excite the $|m|=2$ mode. However, again due to the imperfect excitation, enough of the $|m|=1$ mode was excited to drive the sample into strong nonlinear behavior and saturation before the higher $|m|=2$ mode threshold could be reached.

Finally, it should be mentioned here that the growth constant was found to be an extremely sensitive function of the angle between B_0 and E_0. Figure 14 illustrates this behavior. The value of k_i falls off so rapidly that at an angle greater than $\pm 3°$ the strong attenuation made measurements impossible. The reason for such a striking sensitivity to a very small transverse component of magnetic field is not known. In terms of the physi-

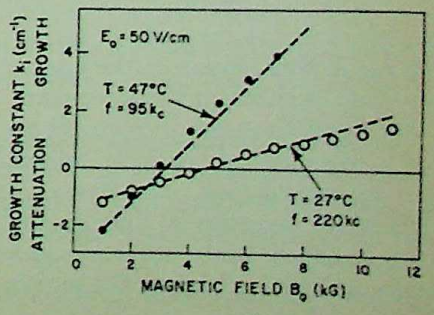


FIG. 12. Growth constant k_i versus B_0 for square sample. Points are experimental; curves are theoretical, calculated from Eq. (32) with $a=0.053$ cm.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

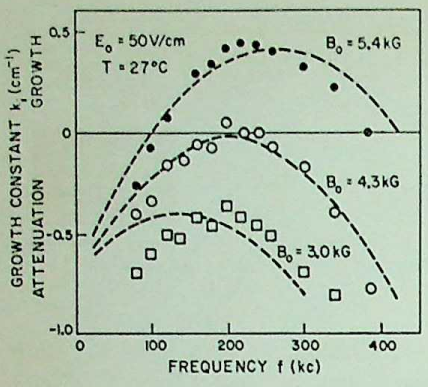


FIG. 13. Growth constant k_i versus f for $B_0 < B_{0c}$, $B_0 \approx B_{0c}$, and $B_0 > B_{0c}$, for square sample. Points are experimental; curves are theoretical, calculated from Eq. (32) with $a = 0.053$ cm.

cal model it would seem that the only effect of such a field is to exert a weak transverse force on the carriers, slightly altering their distribution and setting up a small Hall field. Mathematically, the introduction of a transverse component of B_0 complicates the problem considerably by destroying the azimuthal symmetry. It might be possible to effect some simplification by introducing the transverse component of field as a perturbation. However, such a calculation has not yet been attempted and the effect remains unexplained.

Nonlinear Effects and Current Oscillations

As stated earlier, when signals at the probes reached approximately 500 mV peak-to-peak, saturation began to set in, indicating the onset of nonlinear behavior. The probes farthest down the bar (in the direction of E_0) saturated first as expected, since the amplified signal is largest there. Further increase of the signal level by raising of the excitation or growth rate successively saturated the level at earlier probes, with the signal reaching a limiting value of about 1.5 V peak-to-peak at all probes. At temperatures of 47°C and above, the gain for large enough magnetic fields could be made sufficient to saturate at least the last probes of the sample on amplified internal noise. The spectral characteristics of this noise were quite unusual and will be the subject of further discussion in connection with current oscillations.

Two important effects were noted in conjunction with the onset of strong saturation, namely a marked increase in the dc sample resistance and appearance of oscillations in the terminal sample current. In Fig. 15 the measured dc resistance of the sample is plotted as a function of increasing magnetic field (and hence increasing growth rate). The point at which the resistance begins to increase over its equilibrium value coincides with nearly full saturation at the farthest set of probes (ff′). Current oscillations, shown in the inset of Fig. 15 begin to come into evidence at the same point.

The sample resistance change is a graphic manifestation of the onset of nonlinear behavior. The perturbed carrier density n_1 has now become comparable with the equilibrium densities n_0 and p_0, and the carriers are strongly bunched into the traveling helix. Although the total number of carriers in the sample has not changed, this bunching has produced a net increase in the resistance. The expected order of magnitude of this resistance change may easily be calculated. When strongly bunched, the carriers will be concentrated near the surface of the sample in a helix with pitch $2\pi/k$ and an effective radius approximately equal to the sample radius a. Most of the current will flow between the terminals in this helical path. Both the increased path length of the current and the reduced component of the electric field in the direction of current flow then act to increase the total sample resistance over the equilibrium value. It is a simple matter to show geometrically that the path length is increased by a factor of approximately $(1+k^2a^2)^{1/2}$, and that the component of electric field along the helix is also smaller than the applied fied by the same factor. Thus, with the onset of strong bunching the sample resistance may be expected to increase until it is about a factor of $(1+k^2a^2)$ larger than the unperturbed value. From (26) $k^2a^2 \approx 4/3$, so the resistance should approach a value about 2.3 times larger than that for no bunching. The measured values shown in Fig. 15 are in qualitative agreement with this estimate.

On the basis of the linear theory no terminal VI oscillations are predicted. Since all time-varying quantities vary as $\exp[i\omega t - ikz - im\phi]$, $m \neq 0$, the density integrated over a cross section is independent of time and hence there can be no oscillation of the total resistance. Glicksman's assertion that standing waves, set up by reflections at the contacts, can give rise to terminal oscillations[2] is clearly incorrect since the density still varies sinusoidally with ϕ and t and the integrated density again vanishes. One must therefore invoke nonlinearities or nonuniformities as the basis of the observed oscillation. Misawa[8] has proposed that mixing in of the $m=0$ mode due to deviations from circular symmetry, such as nonuniformities or transverse components of magnetic field, could account for the behavior, and this possibility cannot be ruled out. In fact Okamoto *et al.*[6] have observed that certain oscillistor samples show current oscillations only in the presence of a small transverse component of magnetic field. However, the above experimental observations that oscillations are present in extremely uniform cylindrical samples with strictly longitudinal fields, and that the onset of current oscillation is always associated directly with the onset of nonlinear behavior, provide

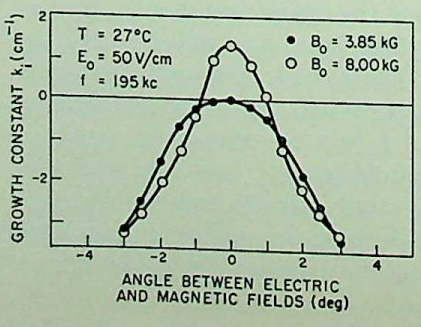


FIG. 14. Growth constant k_i versus angle between E_0 and B_0.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

quite conclusive evidence that large-signal nonlinearity is in itself sufficient to account for terminal oscillations.

The fact that current oscillations were observed in the region of *stable* amplification, requiring only that the gain and excitation be sufficiently large to reach saturation, shows further that instability is not required to obtain such oscillations. In this region the oscillations were essentially sinusoidal at the frequency of excitation, until very strong saturation occurred and some higher harmonic content became apparent. At temperatures below 67°C when the gain was raised sufficiently high to saturate the sample on its own noise, the resultant small current oscillations exhibited a similarly noisy frequency spectrum. However, at 67°C and above, further increase in the magnetic field resulted in a rapid narrowing of the frequency spectrum of both the signals at the probes and the current oscillations, until the frequency was nearly monochromatic and roughly equal to the threshold frequency. One might conceive of a situation in which the strong excursion into nonlinearity might produce a substantial narrowing of the bandwidth for amplification; however, it seems more satisfying to picture the cause as the onset of an instability and the consequent buildup of oscillation at a specific frequency. Indeed just such an instability was predicted from the theory for nearly intrinsic material (i.e., at high temperatures).

Unfortunately, as discussed earlier in connection with the instability theory, a region of stable growth precedes the onset of instability in extrinsic material. The resulting amplification of noise in anything but very nearly or absolutely intrinsic material drives the sample into saturation before instability is attained, so that the linear theory is no longer valid. Such was the case in all of our measurements, even at the highest temperature (87°C). We may, however, make the following qualitative observation. For the 30 ohm-cm germanium used in the measurements, and for $E_0=60$ V/cm and $B_0=11$ kG, the maximum values available with the present arrangement, Eq. (42) predicts that instability will be attainable for temperatures above 60°C. We note also that the rapid narrowing of the frequency spectrum of the amplified noise with increasing fields was observed experimentally only at temperatures above about 67°C. These two facts lead one to the plausible conclusion that this spectral narrowing is the result of the onset of instability and the resulting buildup of oscillations in a narrow band of frequencies; however, further quantitative verification of the theory would certainly be desirable.

Just how close to intrinsic the material must be to reach the instability threshold before nonlinearity occurs may be seen by returning to Eq. (37a) and solving for the growth constant k_{0i} at the point of instability. Neglecting the term $\mu_e\mu_h B_0^2$ which will almost always be negligible, the result is

$$k_{oi}\approx(\mu_a E_0/4D_a)=(qE_0/4kT)[(n_0-p_0)/(n_0+p_0)].$$

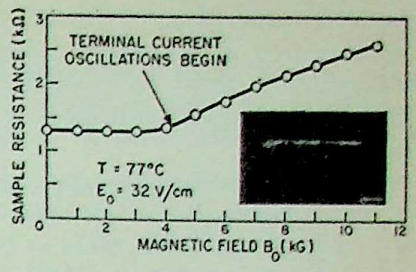


FIG. 15. Average sample resistance versus B_0. Inset is oscilloscope trace of terminal current oscillations which begin at the point indicated. Vertical scale: 10 mA/div. Horizontal scale: 200 μsec/div.

Experimentally it was observed that saturation began to set in when $k_i\approx 4$ cm^{-1}, and in order to keep $\mu_{e,h}B_0\ll 1$ so that the theory correctly applies, the minimum value of E_0 for instability is about 25 V/cm. Therefore we need to have

$$(q/kT)|(n_0-p_0)/(n_0+p_0)|<0.64\text{ V},$$

which for 30 ohm-cm germanium requires $T>100$°C. An easier method would be to start with much purer material. For $N_D-N_A<10^{12}$ cm^{-3} the threshold could be reached while still in the linear region at room temperature.

The Oscillistor

The above prediction and probable observation of absolute instability in a semiconductor bar, with nearly equal densities of electrons and holes and subjected to sufficiently large parallel and magnetic fields, leads naturally to a discussion of the oscillistor, for it is under precisely these conditions that the effect is observed.[3–5] It was shown experimentally that large signal nonlinearity could give rise to terminal VI oscillations and, of course, the instability can provide the mechanism for the rapid buildup of large signal behavior in the form of a strongly bunched helix of carriers.

Unfortunately the oscillistor effect as it is usually observed and reported in the literature presents such a complicated experimental situation that quantitative comparison with theory is prevented. The attainment of nearly equal carrier densities by double injection into a relatively pure bar of semiconducting (long *p-i-n* structure) leads not only to radial gradients of the carrier distribution but also to substantial longitudinal variations in the steady-state densities and electric field.[39] Such variations are not taken into account by the theory and to do so would hopelessly complicate the mathematics. Further, oscillistor samples are usually made with relatively low surface recombination, and as a result growth is probably due both to the steady-state density gradient and to some piling up of carriers at the surface.

Notwithstanding these complications, the essential features of the instability mechanism remain the same, and one would predict the occurrence of an absolute helical instability under the conditions pertinent to the

[39] M. A. Lampert and A. Rose, Phys. Rev. 121, 26 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

oscillistor effect. This result coupled with experimental confirmation of the helical nature of the effect,[6,12,13] provides conclusive verification of Glicksman's proposal of a helical instability as the basis of the phenomenon.

Finally, it should be pointed out that for the gaseous plasma the unperturbed densities of electrons and ions are always equal and there is no ambipolar drift of the helical perturbation. There will therefore be no region of stable growth and, as assumed by Kadomtsev and Nedospasov[1] and Hoh and Lehnert,[21,22] only helical instability is possible.

VI. CONCLUSION

It is evident from the excellent agreement between experiment and theory that the treatment of Sec. III provides an accurate description of the properties of helical waves in the approximation $\mu^2B^2\ll1$. Further, all observations, except the unexpectedly strong dependence of the growth rate on the angle between the electric and magnetic fields, are consistent with the physical model presented in Sec. II.

Although the work was devoted almost entirely to the surface-density mode of operation, it is clear that the general arguments and results may be carried over to more complicated situations, such as the equilibrium-gradient mode or the case of an injected plasma. In addition, although germanium was the only material studied, other semiconductors and perhaps semimetals should exhibit growing helical waves for wide ranges of fields and frequencies, and in fact, oscillistor behavior has been observed in silicon and indium antimonide.[5,10] Unfortunately, in most materials surface and bulk recombination will produce such strong damping of the waves that growth may be difficult to achieve.

ACKNOWLEDGMENTS

We wish to express our appreciation to A. Bers for many fruitful discussions regarding the analysis of gain and instability. We are also grateful to D. K. Hartman of the General Electric Company for supplying several germanium crystals, A. J. Strauss for help with material evaluation, J. H. R. Ward for valuable technical assistance, and Mrs. Nancy Rawson for performing the numerical computations.

APPENDIX A: THRESHOLD CONDITIONS FOR EQUILIBRIUM GRADIENT MODE

We again consider a cylindrical sample of radius a, long in the z direction, but now assume that the equilibrium carrier densities n_0 and p_0 are graded in the radial direction. These gradients could be produced, for example, by the outdiffusion of impurities (consider donors to be specific) in which case n_0 and p_0 would be of the form[40]

$$\begin{aligned} n_0(r)&=n_0J_0(\alpha r)\,,\\ p_0(r)&=n_i^2/n_0(r)=p_0/J_0(\alpha r)\,, \end{aligned} \tag{A.1}$$

where n_i is the intrinsic density. For simplicity we consider the case of weak gradients ($\alpha^2a^2<2$); n_0 and p_0 may then be written as

$$\begin{aligned} n_0(r)&\approx n_0(1-\alpha^2r^2/4)\\ p_0(r)&\approx p_0(1+\alpha^2r^2/4)\,. \end{aligned} \tag{A.2}$$

We further assume that the surface recombination velocity s is large ($s=\infty$), in order to avoid considering an admixture of the equilibrium-gradient and surface-density modes. As usual it is assumed that the bulk lifetime $\tau_b\gg a^2/D_a$ and that the squares of the Hall angles $\mu_{e,h}^2B_0^2\ll1$.

In equilibrium the continuity equations (6) and (7) become, denoting equilibrium values by the subscript 0,

$$\begin{aligned} n_0(r)\mu_e\nabla\cdot\mathbf{E}_0+\mu_e\mathbf{E}_0\cdot\nabla n_0(r)+D_e\nabla^2n_0(r)\\ +\mu_e^2\mathbf{E}_0\cdot[\nabla n_0(r)\times\mathbf{B}_0]=0\,, \end{aligned} \tag{A.3}$$

$$\begin{aligned} p_0(r)\mu_h\nabla\cdot\mathbf{E}_0+\mu_h\mathbf{E}_0\cdot\nabla p_0(r)-D_h\nabla^2p_0(r)\\ -\mu_h^2\mathbf{E}_0\cdot[\nabla p_0(r)\times\mathbf{B}_0]=0\,. \end{aligned} \tag{A.4}$$

The magnetic field B_0 as before is in the z direction, whereas the electric field E_0, in addition to the applied z component E_{0z}, now has a radial ambipolar component E_{0r} due to the density gradients. With equilibrium densities as given in (A.2), Eqs. (A.3) and (A.4) may be readily solved for E_{0r}, with the result (assuming $T_e=T_h=T$)

$$E_{0r}\approx(\alpha^2kT/2q)r\,. \tag{A.5}$$

Following the solution for the surface-density mode, the continuity equations are linearized about equilibrium and quasineutrality is assumed. We then have

$$\begin{aligned} \frac{\partial n_1}{\partial t}&=-\mu_en_0(r)\nabla^2\psi_1+\mu_en_1\frac{1}{r}\frac{\partial}{\partial r}(rE_{0r})\\ &+\mu_e\left(E_{0r}\frac{\partial n_1}{\partial r}+E_{0z}\frac{\partial n_1}{\partial z}\right)-\mu_e\frac{\partial n_0(r)}{\partial r}\frac{\partial\psi_1}{\partial r}+D_e\nabla^2n_1\\ &+\mu_e^2B_0E_{0r}\frac{1}{r}\frac{\partial n_1}{\partial\phi}+\mu_e^2B_0\frac{1}{r}\frac{\partial\psi_1}{\partial\phi}\frac{\partial n_0(r)}{\partial r}=0\,, \end{aligned} \tag{A.6}$$

and

$$\begin{aligned} \frac{\partial n_1}{\partial t}&=\mu_hp_0(r)\nabla^2\psi_1-\mu_hn_1\frac{1}{r}\frac{\partial}{\partial r}(rE_{0r})\\ &-\mu_h\left(E_{0r}\frac{\partial n_1}{\partial r}+E_{0z}\frac{\partial n_1}{\partial z}\right)+\mu_h\frac{\partial p_0(r)}{\partial r}\frac{\partial\psi_1}{\partial r}+D_h\nabla^2n_1\\ &+\mu_h^2B_0E_{0r}\frac{1}{r}\frac{\partial n_1}{\partial\phi}+\mu_h^2B_0\frac{1}{r}\frac{\partial\psi_1}{\partial\phi}\frac{\partial p_0(r)}{\partial r}=0\,. \end{aligned} \tag{A.7}$$

[40] The problem is analogous to that of heat flow from a cylindrical rod. See H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Oxford University Press, London, 1959), Ch. VII.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Due to the presence of terms arising from the equilibrium gradients, the equation for n_1 may not be decoupled as in the surface-mode solution (9). The coupled equations (A.6) and (A.7) are difficult to treat so, following the Kadomtsev and Nedospasov[1] treatment of the very similar gas-plasma case, we solve the problem only approximately in the following manner. The perturbations are Fourier analyzed as in Eq. (11) and only the lowest $|m|=1$ mode is considered. It is now assumed that $N_1(r)$ and $\Psi_1(r)$ have the approximate form

$$N_1(r)=c_1J_1(\beta_1 r),\qquad \Psi_1(r)=c_2J_1(\beta_1 r), \tag{A.8}$$

where $\beta_1 a$ is the first zero of J_1, equal to 3.83, and the c's are constants. Note that $n_1(a)=0$, as required by the assumption $s=\infty$. Equations (A.2), (A.5), and (A.8) are substituted into (A.6) and (A.7) and the resulting equations are averaged over r by multiplying them by $rJ_1(\beta_1 r)$ and integrating with respect to r.[41]

We thus obtain the following two algebraic equations for c_1 and c_2:

$$c_1\{2D_e(k^2+\beta_1^2)-\alpha^2(kT/q)(1+R)+i[2(\omega+\mu_eE_{0z}k)+\alpha^2(kT/q)\mu_e^2B_0]\}+c_2n_0\mu_e\{2(k^2+\beta_1^2)(Q-1)+\alpha^2(1-R)-i\mu_eB_0\alpha^2\}=0, \tag{A.9}$$

and

$$c_1\{2D_h(k^2+\beta_1^2)+\alpha^2(kT/q)(1+R)+i[2(\omega-\mu_hE_{0z}k)+\alpha^2(kT/q)\mu_h^2B_0\}+c_2p_0\mu_h\{2(k^2+\beta_1^2)(Q+1)+\alpha^2(1-R)+i\mu_hB_0\alpha^2\}=0, \tag{A.10}$$

where

$$R\equiv\beta_1\left[\int_0^a rJ_0(\beta_1 r)J_1(\beta_1 r)dr\right]\times\left[\int_0^a rJ_1^2(\beta_1 r)dr\right]^{-1}, \tag{A.11}$$

$$Q\equiv(\alpha^2/4)\left[\int_0^a r^3J^2(\beta_1 r)dr\right]\left[\int_0^a rJ_1^2(\beta_1 r)dr\right]^{-1}. \tag{A.12}$$

The integrals (A.11) and (A.12) are readily evaluated,[42] with the result $R=0$ and $Q=\alpha^2/12\beta_1^2$.

Equations (A.9) and (A.10) are a pair of linear homogenous equations in c_1 and c_2. The dispersion relation is obtained by setting the determinant of the equations equal to zero. Under the assumptions $\alpha^2\ll\beta_1^2$ and $\mu_{e,h}^2B_0^2\ll1$, we then have, after some straightforward algebra,

$$2D_a(n_0\mu_e+p_0\mu_h)(\beta_1^2+k^2)^2+\alpha^2E_{0z}B_0k\mu_e\mu_h(n_0\mu_e-p_0\mu_h)-\alpha^2\omega B_0(n_0\mu_e^2+p_0\mu_h^2)+i\alpha^2B_0\mu_e\mu_h(kT/q)(\beta_1^2+k^2)\times(n_0\mu_h+p_0\mu_e)+i\alpha^2B_0(\beta_1^2+k^2)(n_0\mu_e^2D_h+p_0\mu_h^2D_e)+i2(\beta_1^2+k^2)(n_0\mu_e+p_0\mu_h)(\omega-\mu_aE_{0z}k)=0, \tag{A.13}$$

where we have employed the definitions (10) of μ_a and D_a.

At threshold ω and k are real and (A.13) may be separated into real and imaginary parts. Elimination of ω from the resulting two equations yields

$$(\beta_1^2+k^2)^2+\frac{\alpha^2E_{0z}B_0k}{2D_a}\frac{n_0p_0\mu_e\mu_h(\mu_e^2-\mu_h^2)}{(n_0\mu_e+p_0\mu_h)^2}=0. \tag{A.14}$$

Minimizing the product $E_{0z}B_0$ with respect to k we obtain the threshold conditions:

$$k_c^2=\beta_1^2/3=4.89/a^2, \tag{A.15}$$

and

$$B_{0c}=-\frac{32\beta_1^2k_c}{3m\alpha^2}\frac{D_a(n_0\mu_e+p_0\mu_h)^2}{n_0p_0\mu_e\mu_h(\mu_e^2-\mu_h^2)}\frac{1}{E_0}. \tag{A.16}$$

The result, as expected, is very similar to that for the surface-density mode. The fact that B_{0c} appears to become infinite when $\mu_e=\mu_h$ is a consequence of the fact that the equilibrium gradients of n_0 and p_0 are in opposite directions, producing a partial cancellation of the growth mechanism. In the simple approximation (A.2) used here these gradients are exactly equal and opposite, so that the cancellation appears to be complete when $\mu_e=\mu_h$. In an exact solution this cancellation would be substantial but not complete.

The threshold frequency is obtained from (A.13) where, as discussed in the derivation of (28) and (29) for the surface-density mode, we take care to remain consistent with the assumption $\mu_{e,h}^2B_0^2\ll1$. We then have for $|(n_0-p_0)/(n_0+p_0)|\gg\mu_{e,h}^2B_0^2$,

$$f_c=(1/2\pi)\mu_aE_{0z}k_c \tag{A.17}$$

which, as before, corresponds to the translation of the helix with the ambipolar drift velocity in extrinsic material.

For intrinsic material, $|(n_0-p_0)/(n_0+p_0)|\ll\mu_{e,h}^2B_0^2$,

$$f_c=-\frac{\alpha^2B_{0c}}{2\pi}\left[\frac{kT}{q}\frac{\mu_e\mu_h(n_0\mu_h+p_0\mu_e)}{n_0\mu_e+p_0\mu_h}+\frac{n_0\mu_e^2D_h+p_0\mu_h^2D_e}{n_0\mu_e+p_0\mu_h}\right]. \tag{A.18}$$

The first term is the rotation due to the built-in ambipolar field E_{0r} and the second term arises from the diffusion and drift of carriers across the magnetic field. Equation (A.19) may be simplified considerably by using the Einstein relation $D_{e,h}=(kT_{e,h}/q)\mu_{e,h}$, assuming $T_e=T_h=T$, and by setting $n_0=p_0$, since (A.19) applies to the nearly intrinsic case. Then

$$f_c=-\frac{m\alpha^2}{2\pi}\frac{kT}{q}\mu_e\mu_hB_{0c}. \tag{A.19}$$

[41] See Refs. 1 and 22 for discussion of the validity of this approximate method.

[42] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, London, 1958), 2nd ed., pp. 132–138.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX B: EFFECT OF SMALL SURFACE AND BULK RECOMBINATION

Equation (9) with the inclusion of the bulk term is

$$\frac{\partial n_1}{\partial t}=D_a\nabla^2 n_1-\mu_a E_0\frac{\partial n_1}{\partial z}-\frac{n_1}{\tau_b}, \tag{B.1}$$

where γ has been replaced by $(n-n_0)/\tau_b=n_1/\tau_b$. Recombination has been assumed to take place without trapping so that a single bulk lifetime τ_b may be used for both carriers. Such a treatment is fully satisfactory for germanium at and above room temperatures. In the Fourier analysis (11), $n\sim e^{i\omega t}$ so that (B.1) can be written in the form

$$(i\omega+1/\tau_b)n_1=D_a\nabla^2 n_1-\mu_a E_0(\partial n_1/\partial z). \tag{B.2}$$

Bulk recombination may therefore be included simply by replacing ω by $(\omega-i/\tau_b)$ in (24).

To include surface recombination we proceed from (21) exactly as in the derivation of Eq. (25), but now retain the terms involving s. As a result for small bulk and surface recombination, (25) is replaced by

$$3k^2a^2+(\mu_M{}^2E_0B_0ka^2/D_a)(1-sa/D_a)+4(1+sa/3D_a)+3a^2/2D_a\tau_b=0, \tag{B.3}$$

where second-order small terms involving $(sa/D_a)^2$ and $(sa/D_a)(a^2/D_a\tau_b)$ and $\mu_{e,h}{}^2B_0{}^2$ have been dropped. Proceeding exactly as in the derivation of (26) and (27) we have finally at threshold:

$$k_c{}^2=\frac{4}{3a^2}\left(1+\frac{sa}{3D_a}+\frac{3a^2}{8D_a\tau_b}\right), \tag{B.4}$$

$$B_{0c}=-\frac{6k_cD_a}{m\mu_M{}^2}\left(1+\frac{sa}{3D_a}\right)\frac{1}{E_0}, \tag{B.5}$$

and

$$f_c=(1/2\pi)\mu_aE_0k_c, \qquad \left|\frac{n_0-p_0}{n_0+p_0}\right|\gg\mu_{e,h}B_{0c}{}^2 \tag{B.6}$$

$$=-\frac{20mD_a}{9\pi a^2}(\mu_e-\mu_h)B_{0c}\left(1+\frac{2sa}{3D_a}\right),$$

$$\left|\frac{n_0-p_0}{n_0+p_0}\right|\ll\mu_{e,h}{}^2B_{0c}{}^2. \tag{B.7}$$

As expected, recombination, being a loss mechanism, has raised the threshold fields. However, for the experimental situation of interest: $s\approx30$ cm/sec, $\tau_b\approx1400$ μsec, $a=0.058$ cm, and $D_a\approx60$ cm^2/sec, the correction to k_c and f_c is about 1.5% and that to B_{0c} is about 2.5%. These corrections are less than the experimental error and are therefore negligible.

APPENDIX C: THRESHOLD CONDITIONS FOR GENERAL AZIMUTHAL MODE

In the dispersion relation (21) we again expand the modified Bessel functions in Taylor series expansions except that now we consider a general m instead of $m=\pm1$. The case of $m=0$ can be treated and eliminated at the outset. In (21), setting $m=0$ we have simply

$$\alpha\beta kI_0(\beta a)I_0(ka)=0. \tag{C.1}$$

The magnetic field does not enter this expression at all, so there is no threshold. The solutions therefore correspond to damped or evanescent waves and need not be considered further.

For $m\neq0$ modes we keep only the first two terms of the expansions and substitute them in (21). Proceeding as in Sec. III we arrive at the equivalent of (24) for general m:

$$\begin{aligned}k^2a^2[&|m|(|m|+2)+im(|m|+1)(\mu_e-\mu_h)B_0]\\&+i[a^2(\omega-\mu_aE_0k)/2D_a][|m|(|m|+2)\\&+im(|m|+2)\mu_HB_0-im|m|\mu_aB_0]\\&+m\mu_M{}^2E_0B_0ka^2/D_a+2m^2(|m|+1)\\&+2m|m|(|m|+1)(\mu_e-\mu_h)B_0=0.\end{aligned} \tag{C.2}$$

Separating the real and imaginary parts of (C.2) for real k and ω and eliminating ω from the resulting pair of equations, we then have

$$|m|(|m|+2)k^2a^2+(m\mu_M{}^2E_0B_0/D_a)ka^2+2(|m|+1)=0. \tag{C.3}$$

Therefore, at threshold

$$k_c{}^2=2|m|(|m|+1)/(|m|+2)a^2 \tag{C.4}$$

and

$$B_{0c}=-\frac{2m(|m|+2)D_ak}{|m|\mu_M{}^2}\frac{1}{E_0}. \tag{C.5}$$

From (C.5) it is evident that the $m=1$ mode has the lowest threshold for growth.

Returning to (C.2) we have, following the derivation of (28) and (29), the threshold frequency:

$$f_c=(1/2\pi)\mu_aE_0k_c, \qquad \left|\frac{n_0-p_0}{n_0+p_0}\right|\gg\mu_{e,h}{}^2B_0{}^2 \tag{C.6}$$

$$=-\frac{2m(|m|+1)(2|m|+1)D_a}{\pi(|m|+2)a^2}(\mu_e-\mu_h)B_{0c},$$

$$\left|\frac{n_0-p_0}{n_0+p_0}\right|\ll\mu_{e,h}{}^2B_0{}^2. \tag{C.7}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Relationship between Neutron-Produced Prismatic Dislocations and the Change in Volume during Annealing in Copper*

I. G. GREENFIELD† AND J. R. DEMOS‡
The Franklin Institute Laboratories, Philadelphia, Pennsylvania
(Received 20 August 1963; revised manuscript received 27 January 1964)

Two groups of copper single crystals of various orientations were irradiated with neutrons in the BNL reactor between 50°C and 70°C; one group was exposed to 0.53×10^{18} *nvt* (total epicadmium flux) and the other 1.1×10^{18} *nvt* (total epicadmium flux). An irradiated and a nonirradiated crystal of each orientation were isothermally annealed in a specially constructed dilatometer at temperatures ranging from 75°C to 350°C, and the irreversible differential changes in length were measured. In addition, electron microscope studies were made of the size distribution of irradiation-produced prismatic dislocations after several of the annealing treatments. These experiments revealed that (1) a direct relationship exists between the magnitude of shrinkage after annealing treatments and the total neutron exposure, (2) during annealing dislocation loops decrease in diameter until they disappear, (3) the decrease in dislocation density is related to the volume shrinkage, (4) the changes in loop distribution could be predicted by a line-tension-climb equation, and (5) the activation energy for the annealing of irradiation-produced loops lies between 2.08 eV and 2.10 eV. The results indicated that the annealing process in neutron-irradiated copper is not the same as observed in quenched metals although a self-diffusion mechanism appears to be the controlling factor.

INTRODUCTION

IRRADIATION of copper with high-energy particles results in defects that produce an expansion in the lattice. Vook and Wert[1] bombarded copper foils with deuterons of an average energy of 8.5 MeV at 15°K and found that copper expanded; subsequent annealing to about 280°K resulted in a shrinkage curve which is the same form as the resistivity curve determined by other workers.[2,3] Although several experiments concerned with the dilatation of copper[4] were made, a detailed account of the effect of annealing after irradiation with neutrons at reactor temperature of approximately 50°C has not been reported. Experiments show, however, that the physical and mechanical properties of irradiated copper are affected by subsequent annealing[4,5] between 0 and 500°C. Moreover, Coltman, Klabunde, McDonald, and Redman[6] have shown that thermal neutrons as well as fast neutrons produce measurable damage at low temperatures; however, more investigation is necessary to ascertain the relationship between the effect of slow neutron bombardment at pile temperatures (50–70°C) and the recoverable damage.

Transmission electron microscopy investigations of neutron-irradiated copper revealed dark spots in copper which are believed to be prismatic dislocation loops.[7,8] These prismatic dislocation loops do not form in all cases, but their production is found to be dependent upon the material irradiated and the temperature of irradiation.[9,10] In copper the density and size distribution of these spots are related to the total radiation exposure and the neutron spectrum.[11] As the irradiation exposure is increased, the density increases to a saturation value and the loops grow in size; subsequent annealing of this material will lower the density of the loops.[7,11,12] Since the disappearance of these loops is equivalent to a decrease in dislocation density, the present experiments were designed to obtain data from which the change in volume of the copper could be correlated with the change in dislocation density during the annealing of irradiated copper. Conclusions can be made concerning (1) the amount of point defects remaining in the lattice after irradiation, (2) the volume occupied by the dislocation loops, and (3) the mode of diffusion during annealing treatments.

EXPERIMENTAL PROCEDURE

Single crystals, 3 mm in diameter by 140 mm long, of high-purity copper (99.9995%) were grown in graphite molds in a vacuum. These crystals were mechanically cut into two equal parts, 70 mm long, in order to make pairs of crystals of identical orientations suitable for the precise length change measurements.

* This research was supported by the U. S. Atomic Energy Commission.

† Present address: University of Delaware, Newark, Delaware.

‡ Present address: U. S. Department of Commerce, National Bureau of Standards, Washington, D. C.

[1] R. Vook and C. Wert, Phys. Rev. **109**, 1529 (1958).

[2] R. O. Simmons and R. W. Balluffi, J. Appl. Phys. **30**, 1249 (1959).

[3] H. G. Cooper, J. S. Koehler, and J. W. Marx, Phys. Rev. **97**, 599 (1955).

[4] For a review, see D. S. Billington and J. H. Crawford, Jr., *Radiation Damage in Solids* (Princeton University Press, Princeton, New Jersey, 1961), Chap. 5, p. 100.

[5] T. H. Blewitt, R. R. Coltman, R. E. Jamison, and J. K. Redman, J. Nucl. Mater. **2**, 272 (1960).

[6] R. R. Coltman, C. E. Klabunde, D. L. McDonald, and J. K. Redman, J. Appl. Phys. **33**, 3509 (1962).

[7] J. Silcox and P. B. Hirsch, Phil. Mag. **4**, 1356 (1959).

[8] I. G. Greenfield and H. G. F. Wilsdorf, Naturwiss. **47**, 395 (1960).

[9] H. G. F. Wilsdorf and D. Kuhlmann-Wilsdorf, Phys. Rev. Letters **3**, 170 (1959).

[10] I. G. Greenfield and H. G. F. Wilsdorf, Conf. J. Phys. Soc. Japan **18**, Suppl. III, 20 (1963).

[11] M. J. Makin, A. D. Whapham, and F. J. Minter, Phil. Mag. **7**, 285 (1962).

[12] M. J. Makin and S. A. Manthorpe, Phil. Mag. **8**, 1725 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Neutron spectrum for irradiation treatments (temperature between 50–75°C).

Exposure	Specimens	Thermal flux	Epicadmium flux	Flux above 1 MeV	Effective time	Total thermal flux	Total epicadmium flux	Total flux 1 MeV
		(10^{13} *nv*)	(10^{11} *nv*)	(10^{11} *nv*)	(10^6 sec)	(10^{19} *nvt*)	(10^{18} *nvt*)	(10^{17} *nvt*)
A	70, 81, 82	1.0	3.8	2.7	3.0	3.0	1.14	8.10
B	66	1.3	2.8	...	1.9	2.47	0.532	...

To reduce the possibility of spurious effects due to oxygen entering the lattice during the initial annealing treatments, the copper specimens were heated in an evacuated quartz tube in the presence of a zirconium getter. After annealing for 12 h one crystal from each orientation was sent to Brookhaven and irradiated in the BNL graphite reactor. Specimens 70, 81, and 82 were irradiated together in the same hole with a total integrated epicadmium flux of 1.14×10^{18} (*nvt*), exposure A, whereas specimen 66 was exposed to about one-half of this dose, exposure B. Table I contains the pertinent data concerning the neutron bombardment.

The dilatometer[13] used in these experiments was designed to compare irreversible differential changes in length resulting from the rearrangement of point defects between an irradiated and a nonirradiated crystal of the same orientation. The dilatometer was similar in construction to the one used by J. Takamura in the study of changes in length of quenched metals after annealing at temperature near 20°C.[14] The present dilatometer can be operated in the temperature range from 75 to 400°C. The relative changes in length between the irradiated and nonirradiated crystals were magnified by an optical level system and a telescope; when necessary, the readings thus procured were recorded by a slow motion cine technique. In order to obtain precision measurements, the dilatometer was kept in a room in which the temperature did not vary more than 1.0°C over a 24-h period. It was necessary, however, to monitor the temperature of certain components of the dilatometer in order that accurate corrections could be applied to the dilatometer scale readings. An analysis of the calibration data has shown that the sensitivity was about 200 Å with the standard deviation of ±500 Å for the maximum variation of temperature of the components of the dilatometer. All annealing experiments were carried out in an atmosphere of high-purity argon that was passed over hot zirconium chips.

The size distribution and the density of dislocation loops were measured from micrographs in regions about 2000 Å thick. The thickness was measured, where possible, by calculating the perpendicular projection to the surface of the intersections of slip traces. Although dislocations producing these traces were active after long exposures to the electron beam, the dislocations were not always present in the regions where the density measurements were made. As a result some interpolation was necessary; however, the accuracy falls within ±500 Å of the thickness measurement.

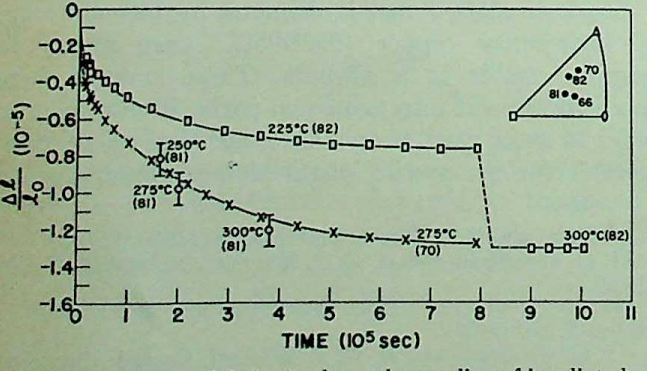


FIG. 1. Change in $\Delta l/l_0$ for isothermal annealing of irradiated copper crystals 70, 81, 82; irradiation exposure A.

[13] I. G. Greenfield and J. Takamura (to be published).
[14] J. Takamura, Acta Met. 9, 547 (1951).

RESULTS

Length Changes Due to Annealing

The effect of isothermal annealing on the change in length of single crystals of copper irradiated with neutron exposure A is shown in Fig. 1. In this graph of specific length change, $\Delta l/l_0$, is plotted against annealing times at the desired temperature. Consider first the results of annealing at 225°C (specimen 82) for 7.9×10^5 sec. The continuous record of the length changes has shown a high rate of shrinkage during the early part of the annealing treatment; however, after approximately 6×10^5 sec, the rate of shrinkage became negligible. Precise measurements of the irreversible difference in length were then made at room temperature after the irradiated specimen and the dummy specimen were cooled to room temperature; the measurements agreed with those taken at the annealing temperature. Then the specimen was heated to 300°C and again a rapid shrinkage took place during the initial heating; however, after about 0.5×10^5 sec, $\Delta l/l_0$ became constant. The total value for the specific change in length for 225°C was found to be -0.76×10^{-5} and for 300°C, -1.30×10^{-5}.

The isothermal annealing treatment carried out at 275°C (specimen 70) was also characterized by an initial rapid shrinkage rate that was greater than observed with the 225°C annealing treatment, and after about 7×10^5 sec, $\Delta l/l_0$ reached a constant value. The total specific length change was not significantly

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

different from the value of -1.30×10^{-5} found in the 300°C annealing treatment discussed above. The irradiated crystal treated at 275°C was not annealed at a higher temperature since it was necessary to preserve the prismatic-dislocation-loop distribution in the specimen for an electron microscope investigation that will be discussed later in this paper. Specimen 81 was annealed at 250, 275, and 300°C for lengths of time indicated in Fig. 1. The results of these dilatometric measurements agree with those described earlier. Annealing at higher temperatures, up to 350°C, did not produce significant additional shrinkage, and thus the maximum $\Delta l/l_0$ for this neutron exposure was -1.30×10^{-5}.

Thus, for the neutron dose A no detectable shrinkage occurred below 200°C; above 225°C considerable shrinkage initially took place, and, as the annealing time was increased, the shrinkage rate approached zero. Additional shrinkage occurred when the temperature was raised higher than this annealing temperature; however, after being annealed in the range of 275°C for more than 7×10^5 sec or 300°C for a shorter time, no additional changes in $\Delta l/l_0$ are detected. The maximum $\Delta l/l_0$ was about equal to -1.30×10^{-5} for this irradiation dose.

For the lower irradiation exposure B, on the other hand, the shrinkage behavior was somewhat different. Specimen 66 was annealed at temperatures and times listed in Table II. The annealing treatment was extended until it was certain the rate of change of $\Delta l/l_0$ reached zero; then the specimens were cooled to room temperature before the next isothermal treatment. The times listed in Table II are not necessarily the minimum time for $\Delta l/l_0$ to reach a constant value. Above 325°C additional shrinkage was not seen. If the specific shrinkages in Table II are compared with those shown in Fig. 1, it can be concluded that the cumulative $\Delta l/l_0$ for the crystals exposed to the smaller neutron dose approached a constant value at about 325°C, whereas for the greater dose the maximum $\Delta l/l_0$ is found at about 275°C. In addition, the maximum shrinkage for the lower dose is less than one-half that for the higher irradiation exposure. Thus, the range of temperature in which shrinkage of the specimen occurs and the total $\Delta l/l_0$ attainable depends upon the irradiation dose.

TABLE II. Annealing treatments for specimen 66; neutron exposure B.

Annealing temperature (°C)	Time at temperature (10^5 sec)	Cumulative $\Delta l/l_0$ (10^{-5})
75, 100, 150, 175, 200	...	...
250	2.52	0.24
275	1.80	0.34
300	3.42	0.75
325	2.16	0.82
350		0.82

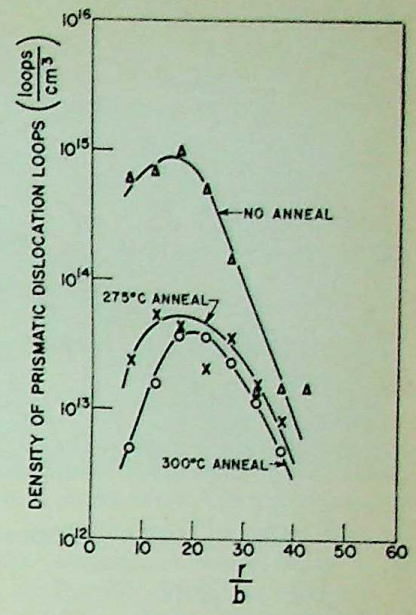


FIG. 2. Size distribution of prismatic dislocation loops in neutron irradiated copper after annealing.

Density and Distribution of Dislocation Loops after Annealing

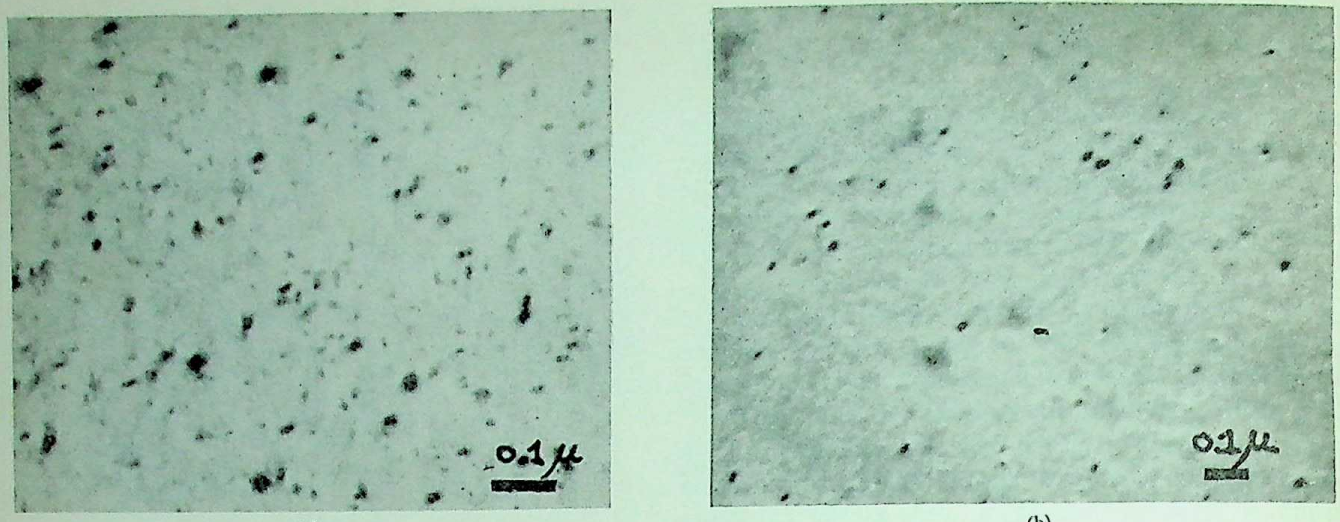
The distributions of prismatic dislocation loops in neutron irradiated copper before and after annealing are shown in Fig. 2. The parameter r/b is the radius of the loop divided by the distance of closest approach between atoms. The upper curve represents the configuration observed in the copper crystals following irradiation treatment A (see Table I). Figures 3(a) and 3(b) are electron micrographs of specimens in the as-irradiated state and the annealed state. Small dark spots less than 25 Å in diameter were not seen in abundance as reported by Makin *et al.*[11] The geometry of the diffraction contrast in the small spots did not reveal, in general, the nature of the irradiation-produced defects, and consequently, it was not possible to infer whether these structures were tetrahedrons or small dislocation loops. Since the density and size distribution were measured in rather thick regions, the amount of inelastic scattering of the electrons has contributed to the limited image resolution.

The density of dislocation loops were decreased more than a factor of 10 after annealing crystal 70 at 275°C see Fig. 2. Moreover, the average size of loops has only increased slightly, indicating that loop growth plays a minor role during the annealing of neutron irradiated copper. After treating specimen 82 at 300°C, the loop distribution was not altered greatly although the average loop radius was slightly higher. Figure 3 illustrates the typical micrographs before and after annealing. It should be pointed out that loop size measurements are most accurate at the large-radii part of the distribution curve, and the density measurements are most accurate when the loop density is above 10^{13}/cc.

The most significant result due to annealing is the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri



(a) (b)

FIG. 3. Electron micrographs of irradiated copper (exposure A) (a) before annealing and (b) after annealing at 275°C for 7×10^5 sec.

great decrease in density of dislocation loops with only a slight increase in average radii of the loops.

ANALYSIS OF RESULTS

Dependence of Prismatic Loop Size on Annealing Conditions

Silcox and Whelan[15] have demonstrated that the shrinkage of prismatic dislocation loops formed by quenching in aluminum can be described by an equation originally derived by Friedel.[16] It was shown that the rate of change of the radius of a prismatic-dislocation loop is related to three temperature-dependent conditions: (1) diffusivity of the point defects, (2) the line tension of the dislocation loop and (3) the supersaturation of point defects in the vicinity of the loop.

The rate of change of the radius is then

$$dr/dt=-Z\nu_a b\exp(-E/kT)\times[\exp(F_c b^2/kT)-\exp(F_s b^2/kT)]C_c, \quad (1)$$

where Z is the atomic coordination number, ν_a is the atomic frequency, b is the Burgers vector, E is the activation energy for vacancy diffusion, $F_c=Gb\ln(r/b)/[4(1-\sigma)r/b]$, G is the shear modulus, σ is Poisson's ratio, and $F_s=(kT/b^2)\ln(c/c_0)$. The factor $\exp(F_c b^2/kT)$ is dependent upon the line tension of the loop, and the factor $\exp(F_s b^2/kT)$ is the supersaturation term which indicates whether the loop will accept or emit vacancies. The term C_c is considered to be the concentration of jogs and in a prismatic dislocation loop can be considered as equal to $\frac{1}{2}$. Friedel[17] has pointed out that this analysis depends upon the assumption that the growth process is controlled by climb and not by the migration of point defects to sinks. It has been shown that when loops are considered in thin foils, Eq. (1) can be simplified by considering the surface as a sink for the vacancies. Then c/c_0 is approximately equal to unity (i.e., nearly an equilibrium concentration of vacancies), and Eq. (1) can be written as

$$dr/dt=-\tfrac{1}{2}Zb\nu_a\exp(-E/kT)[\exp(F_c b^2/kT)-1]. \quad (2)$$

Since prismatic dislocation loops are produced in copper by neutron bombardment, it may be assumed (1) that shrinkage of these irradiation-produced loops is controlled by the same forces as the quenched-in loops, and (2) sinks and sources of point defects are available during annealing, and (3) the average distance between sinks and sources of point defects is about the same as the average distance between loops. In contrast to the annealing behavior of isolated loops which have been studied,[15] this investigation is concerned with a multitude of small loops in bulk material, however, if the assumption is made that the vacancies move to sinks faster than they are produced by the loops and clusters, then Eq. (2) can be used.

For aluminum and copper which contain loops of sizes larger than 50 atomic diameters in radius, this equation may be simplified since the reciprocal of the bracketed term is merely a linear function of r/b.

For loops of smaller radii, this relationship is not linear; and, to determine the time required for a prismatic dislocation loop to shrink to a given size, it is necessary to consider the bracketed term in Eq. (2) for various values of r/b (radius/Burgers vector). If Eq. (2) is rearranged into the following form and integrated between the limits indicated,

$$\int_{t_0}^{t}dt=-\frac{2}{Z\nu_a}\exp\left(\frac{E}{kT}\right)\times\int_{r_0/b}^{r/b}\left[\exp\left(\frac{F_c b^2}{kT}\right)-1\right]^{-1}d\left(\frac{r}{b}\right), \quad (3)$$

[15] J. Silcox and M. J. Whelan, Phil. Mag. 5, 1 (1960).
[16] J. Friedel, *Les Dislocations* (Gauthier-Villars, Paris, 1956).
[17] J. Friedel's discussion at the end of Ref. 13.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

then the time necessary for a prismatic dislocation loop climbing from radius r_0 to radius r can be determined.

Equation (3) can be integrated numerically after the values of the constants are substituted. It has been found that the appropriate value for the activation energy E falls between 2.08 and 2.10 eV (the evidence for these values will be discussed later) and the relationships between r/b and time for various size loops are shown in Figs. 4 and 5. In Fig. 4 the solid lines represent the theoretical shrinkage of dislocation loops at 275° with a self-diffusion activation energy of 2.08 eV. The initial values for r/b are found at the ordinate when the annealing time is zero and the values of r/b decrease in a homologous manner during the annealing treatment.

From this graph the change in loop sizes can be estimated for the irradiated copper single crystal 70 after annealing at 275°C for 7.9×10^5 sec. For example the loops represented by $r/b=60$ shrink to $r/b=58$; the loops of initial size $r/b=30$ shrink to $r/b=19$; and loops of initial sizes below about $r/b=26$ completely disappear after this annealing treatment.

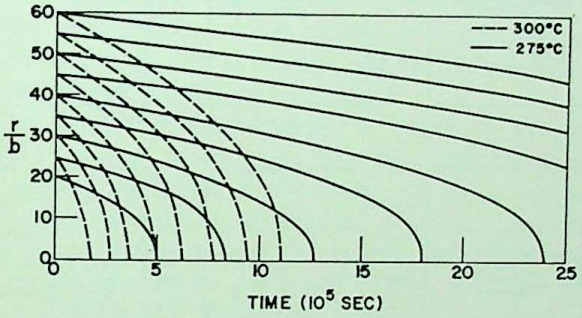


FIG. 4. Shrinkage of loops of size r/b when $E=2.08$ eV at 275 and 300°C.

A similar analysis can be made from the data in Fig. 5 where $E=2.10$ eV. With this information and the original size distribution after irradiation, a new size distribution can be estimated after an annealing treatment; these results are shown in Fig. 6. The shaded area represents the expected values for the density distribution when activation energy between 2.08 and 2.10 eV is used.

For r/b greater than 25, the experimental curve falls in the shaded region of the estimated curves, however, for low values of r/b the estimated curve rises above the experimental curve and ends in a region where $r/b\simeq15$. The differences between the experimental and estimated values of size distribution in the region where $r/b<25$ can be ascribed to the experimental difficulty in positively identifying all the small loops. Thus, the measured density values would tend to be below the theoretically estimated values which were derived from the distributions of larger loops. In addition, the theoretical description of the shrinkage of loops predicts that all loops with initial sizes $r/b<26$ will

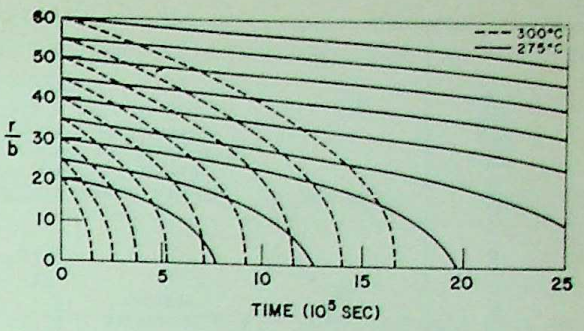


FIG. 5. Shrinkage of loops of size r/b when $E=2.10$ eV at 275 and 300°C.

disappear during the annealing treatment; consequently, the estimated curve results in a sharp cutoff. It is more probable, however, that the curve does not abruptly end, and that a maximum would occur as the measured evidence indicates. Thus, a compromise between the measured data and the calculated data is reasonable. If the same analysis is made for specimen 82 (first annealed at 225°C and then annealed at 300°C), a new distribution can be estimated which is shown in Fig. 7. As in the previous case, the new distribution has the proper form and is a reasonable fit with the actual distribution. These estimated curves of size distribution after annealing are affected greatly by the value of the activation energy E, and the best fit has been found to be between 2.08 and 2.10 eV. Other investigators have found activation energies of this magnitude for annealing irradiated copper in this

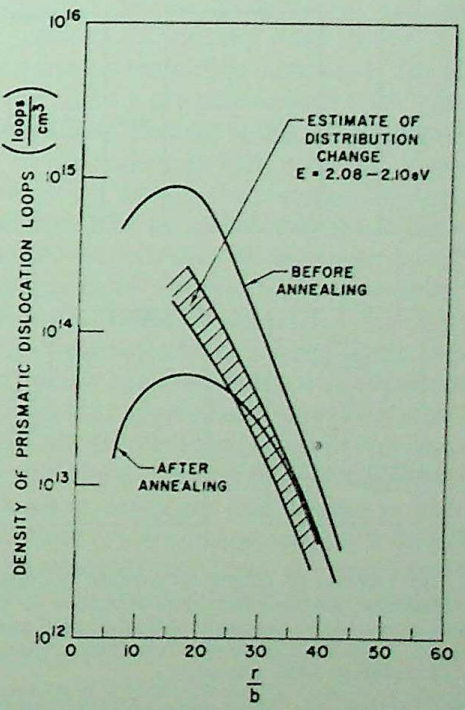


FIG. 6. Estimate of distribution change for annealing treatment at 275°C for 7.9×10^5 sec when 2.08 eV $\leqq E \leqq$ 2.10 eV.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

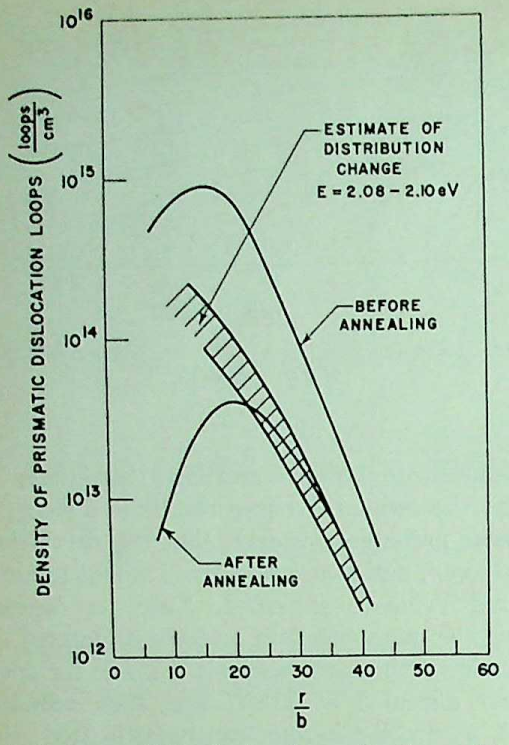


FIG. 7. Estimate of distribution change for the combination of annealing treatments (1) at 225°C for 7.9×10^5 sec and (2) at 300°C for 2.1×10^5 sec when $2.08\ \text{eV} \leqq E \leqq 2.10$.

temperature range.[18,19] Makin[12] has concluded from recent experiments that the annealing of small spots is associated with an activation energy of 2.0 eV and the annealing of the large spots is associated with an activation energy of 1.0–1.5 eV. His determination was not based on the line-tension-climb relationship described by Eq. (1).

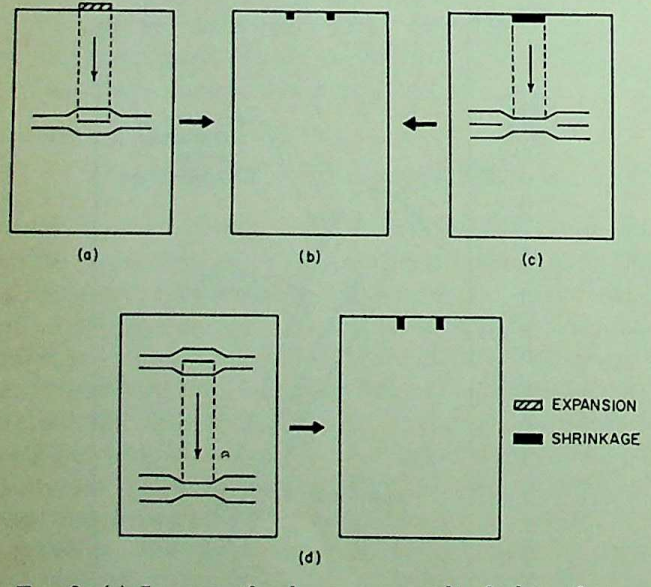


FIG. 8. (a) Increase of volume as a result of plate of atoms moving to surface; (b) shrinkage after relaxation in the vicinity of loop; (c) decrease in volume as a result of plate of atoms moving into loop; (d) transfer of plate of atoms from extrinsic loop and the over-all shrinkage after relaxation.

[18] A. W. McReynolds, W. Augustyniak, M. Mckeown, and D. B. Rosenblatt, Phys. Rev. 98, 418 (1955).
[19] R. R. Eggleston, Acta Met. 1, 679 (1953).

Relationship Between Volume Shrinkage and Dislocation Loops

The specific volume change $\Delta V/V_0$ for isotropic shrinkage is three times the specific length change. Comparisons between the length changes after annealing for the crystals of the orientations shown in Fig. 1 give evidence that $\Delta l/l_0$ is isotropic; hence, the specific volume changes are readily computed.

The relationships between volume change and the removal of dislocation loops from a crystal is illustrated by the schematic diagram in Figs. 8 and 9. In Fig. 8(a) the "plate of atoms" is transported from the dislocation loop to the surface, thus, increasing the volume of the crystal by the extra plate of atoms; a void equal to the

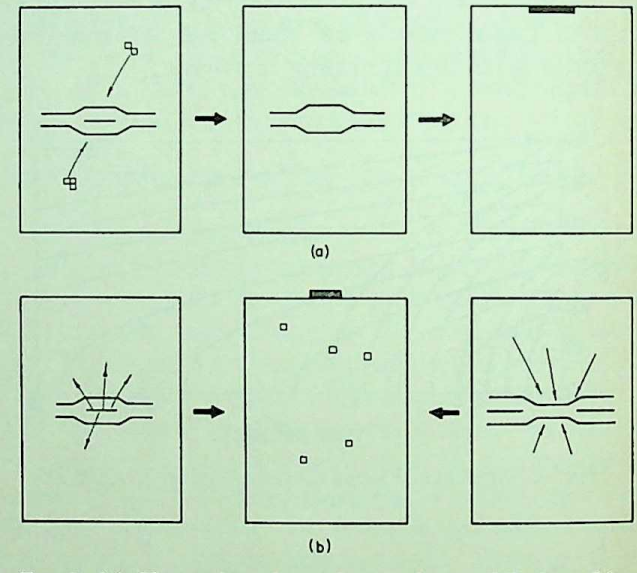


FIG. 9. (a) Change in volume due to the annihilation of interstitial loops by vacancies in the lattice; (b) dissolution of vacancy or interstitial loops.

volume of the atoms transported plus the volume of dislocation of length $2\pi r$ remain in the lattice. After a readjustment of the lattice planes, the differential volume change corresponds to the strain field associated with $2\pi r$ length of dislocation; and, consequently, a small shrinkage occurs [dark areas in Fig. 8(b)]. A similar argument is applicable for a vacancy loop and is diagrammed in Fig. 8(c). Moreover, when both vacancy and interstitial loops co-exist, the "extra plate of atoms" can replace the "plate of vacancies" if the loops are the same size or can replace portions of the vacancy loops if the loops are of different sizes; the decrease in volume is again proportional to the decrease in dislocation length [see Fig. 8(d)]. The same relationship results if loops moved out of the material by slip along the glide cylinder. The volume change due to the annihilation of an interstitial loop by vacancies which are in super-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Specific volume changes based on the reduction of dislocation.

(1)	(2)	(3)	Densities After Annealing (4)	(5)	(6)	(7)	(8)
Annealing temperature °C	Prismatic dislocation density (loops/cm³) 10^{15}	$2\pi\bar{r}$ (10^{-6} cm)	Dislocation density (10^8/cm²)	Total cumulative volume change $\Delta V/V_0$ (10^{-5})	Dislocation density (10^8/cm²)	$\Delta V/V_0$ (1)[a]	$\Delta V/V_0$ (2)[b]
Room	2.90	2.59	75.1	...	...	...	...
275	0.226	3.26	7.37	−3.84	66.8	−0.62	−3.15
300	0.144	3.43	4.94	−3.93	70.2	−0.65	−3.28

[a] The value of $\Delta V/V_0$ (1) is based on the volume associated with the dislocation-density change after annealing.
[b] The value of $\Delta V/V_0$ (2) is based on the summation of the specific volume change in column 7 and the specific volume change associated with the reduction of the supersaturation of point defects.

saturation in the lattice is illustrated in Fig. 9(a). The over-all shrinkage is larger than shown in Fig. 8 and depends upon the area enclosed in the loop, the lattice relaxation around the vacancies or clusters of vacancies, and the volume associated with the dislocation loop. A similar analysis for vacancy loops and a supersaturation of interstitial atoms also indicates volume shrinkage. Another possible interaction between loops and point defects is shown in Fig. 9(b). In this case the loops dissolve into point defects which remain in the crystal, and thus, a volume expansion should be found. Experimentally only shrinkage was measured, and the latter process will not be considered.

The dislocation density, ρ_1, is equal to the total dislocation-loop density, ρ_0, times the circumference of an average loop, $2\pi r$. If the change in ρ_1 is compared with the specific volume change after an isothermal annealing treatment, then the volume for a unit length of dislocation can be evaluated. These data are compiled in Table III. Stehle and Seeger[20] have calculated the change in volume associated with a dislocation by applying a nonlinear model to describe the dilation. Their results indicate that the dilation is greatest in the region from one Burgers vector to the center of the core, and the volume associated with a dislocation of length b is about twice the atomic volume. These calculations were applied to straight dislocations. For dislocation loops, however, the volume per unit length of dislocation loop would be somewhat dependent upon the diameter of the loop. If the loops are considered to be prismatic, and if it is assumed that the volume per unit length is the same as in a straight edge dislocation, the specific volume change associated with the loop density change can be calculated for the various annealing treatments. The results are given in Table III. For the annealing treatments at 275 and 300°C the calculated shrinkage is about one-sixth the measure value [see columns (5) and (7)].

If the decrease in density of the loops is a result of the combination of point defects from defect clusters with the loops as shown in Fig. 9(a), then the specific volume change will be the sum of the volume of the dislocation and the volume previously occupied by the clusters. Assuming that a cluster having n point defects has a volume of $n\times$atomic volume (relaxation not considered), the $\Delta V/V_0$ derived from the experimental data of dislocation density is in good agreement with the experimentally measured values [see columns (5) and (8) in Table III].

The present and other annealing experiments[7,11,12] have demonstrated that the dislocation loops in neutron-irradiated copper will decrease in diameter during annealing until they disappear. A similar loop behavior is noted in quenched metals in regions near sinks such as grain boundaries and dislocations that are able to climb.[15] In contrast, when sinks are not available there is a tendency for loops to grow as evidenced in quenched Al–0.5 at.% Mg.[21] Since the loops decrease in size during annealing in irradiated copper, it is assumed that loops are emitting or accepting defects in such a way as to cause loops to shrink. It is expected that the rate-controlling process is diffusion; and, thus, the average distance a vacancy may move during the annealing treatments can be computed. For the 275°C anneal, the vacancy can travel about 1000 Å in 7.9×10^5 sec; this diffusion distance is too small for vacancies to migrate to the surface in bulk material. However, the dislocation loops are on the average 700 Å apart and can serve as sinks for point defects. A direct determination of the intrinsic or extrinsic nature of the dislocation loops in neutron-irradiated copper has not appeared in the literature although loops produced by α-particle[22] bombardment have been found to be interstitial. A recent electron diffraction experiment concerning loops that appear in molybdenum after neutron irradiation at elevated temperatures has shown conclusively that these loops are interstitial.[23] Makin *et al.*, on the bases of annealing experiments in which large and small spots are seen, have concluded that the large spots are interstitial loops and the small are vacancy clusters.[11,12] The nature of the dislocation loops cannot be determined from the present experiment, but the

[20] H. Stehle and A. Seeger, Z. Physik **146**, 217 (1956).

[21] J. Takamura, K. Okazaki, and I. G. Greenfield, Conf. J. Phys. Soc. Japan **18**, Suppl. III, 78 (1963).
[22] R. S. Barns and D. J. Mazey, Phil. Mag. **5**, 1247 (1960).
[23] J. D. Meakin and I. G. Greenfield (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

evidence indicates that both loops and submicroscopic defect clusters exist in these neutron irradiated specimens.

SUMMARY

Copper crystals irradiated with neutrons between 50 and 70°C show the following characteristics when isothermally annealed between 75 and 350°C:

(1) A decrease in volume is detected and is proportional to the neutron exposure. The maximum specific volume changes for 1.14×10^{18} (*nvt* total epicadmium flux) and 0.532×10^{18} (*nvt* total epicadmium flux) are -3.93×10^{-5} and -2.46×10^{-5}, respectively.

(2) Measurable shrinkage of the single crystals takes place between 225 and 325°C, and the temperature where the maximum shrinkage rate occurs appears to be higher for the lower neutron dose.

(3) The change in size distribution of prismatic dislocation loops can be predicted by Eq. (1).

(4) The activation energy for the annealing process corresponds to the activation energy for self-diffusion.

(5) Dislocation loops and submicroscopic point-defect clusters are present in neutron irradiated copper.

ACKNOWLEDGMENTS

The authors are grateful for the interest and advice of Dr. H. G. F. Wilsdorf of the University of Virginia and Dr. J. Takamura of the University of Kyoto, and for the stimulating discussions with Dr. C. E. Birchenall of the University of Delaware and Dr. R. Vook of The Franklin Institute Laboratories. We are indebted to R. Larson who has taken care of the irradiation of the specimens in the Brookhaven reactor.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Thermal Conductivity of Silicon and Germanium from 3°K to the Melting Point*

C. J. GLASSBRENNER† AND GLEN A. SLACK
General Electric Research Laboratory, Schenectady, New York

(Received 11 December 1963)

The thermal conductivity K of single crystals of silicon has been measured from 3 to 1580°K and of single crystals of germanium from 3 to 1190°K. These measurements have been made using a steady-state, radial heat flow apparatus for $T>300$°K and a steady-state, longitudinal flow apparatus for $T<300$°K to give absolute K values. This radial flow technique eliminates thermal radiation losses at high temperatures. The accuracy of both the low-temperature apparatus and the high-temperature apparatus is approximately ±5%. Some special experimental techniques in using the high-temperature apparatus are briefly considered. At all temperatures the major contribution to K in Si and Ge is produced by phonons. The phonon thermal conductivity has been calculated from a combination of the relaxation times for boundary, isotope, three-phonon, and four-phonon scattering, and was found to agree with the experimental measurements. Above 700°K for Ge and 1000°K for Si an electronic contribution to K occurs, which agrees quite well with the theoretical estimates. At the respective melting points of Si and Ge, electrons and holes are responsible for 40% of the total K and phonons are responsible for 60%. The measured electronic K yields values for the thermal band gap at the melting point of 0.6±0.1 eV for Si and 0.26±0.08 eV for Ge.

INTRODUCTION

IN a semiconductor, various carriers can contribute to the thermal conductivity. These are phonons, photons, electron-hole pairs, and the separate electron and holes. The problem of sorting out the contributions from these processes has been handicapped by the lack of accurate high-temperature measurements. In this paper, we report on measurements of the thermal conductivity K of single crystals of germanium and silicon between 300 and 1580°K taken with an improved cylindrical heat flow apparatus. The measurements of K of silicon below 300°K have been taken with an existing longitudinal heat flow apparatus.[1] Germanium and silicon were measured because they are easily obtained with high purity and known electrical characteristics. As a consequence of this investigation, they could serve as standard materials for K measurements.

The purpose of this investigation was to measure and analyze the K of these materials up to the melting point and, in particular, to unravel the magnitude of the contributions from the various carriers of heat. It was found that the phonon or lattice thermal conductivity is dominant and decreases faster than $1/T$ at high temperatures. The bipolar thermal conductivity, i.e., from electron-hole pairs, is significant at high temperatures and agrees with a simple theory. The polar thermal conductivity, i.e., from the separate electrons and holes, is small, and the photon thermal conductivity is not detectable.

* This work is part of a thesis submitted by C. J. Glassbrenner to the Physics Department of the University of Connecticut in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

† Present address: Controls for Radiation Inc., Cambridge, Massachusetts.

[1] G. A. Slack, Phys. Rev. 122, 1451 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

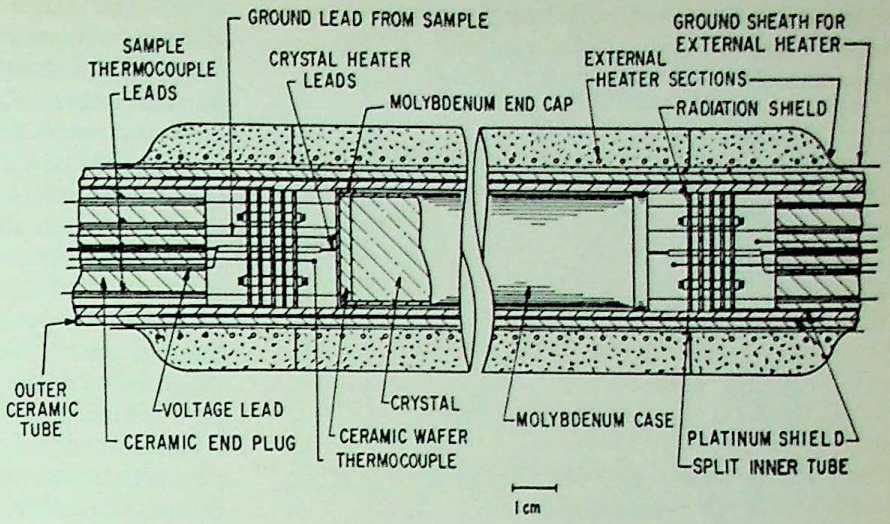


FIG. 1. The high-temperature radial heat flow apparatus showing the crystal sample, heater, thermocouples, and shields.

CRYSTAL SPECIFICATIONS AND PREPARATIONS

A. Low-Temperature Silicon Crystal

The silicon single crystal used for the low-temperature measurements was grown by the pedestal method from a bar of high-purity silicon by W. C. Dash of this laboratory. The growth direction was [111]. It was made with care in order to be oxygen and dislocation free.[2–4] Any vacancy clusters were less than one micron in diameter as determined by etchpit and copper decoration techniques. The pedestal method prevents the silicon from coming in contact with quartz, usually the chief source of oxygen. The crystal was 2 cm long with a geometrical average diameter of 0.44 cm. The diameter varied slightly along the length of the sample. The cross section could be best described as an equilateral triangle with rounded corners, typical of [111] oriented crystals. The room-temperature resistivity was approximately 2000 Ω-cm, and the sample was p type.

B. Crystals Used at High Temperatures

The silicon and germanium samples were both single crystals, cylindrically shaped, approximately 13 cm long by 2.6 cm in diameter. They had on the order of 10^5 dislocations[5] per cm^2. The Ge crystal was grown by the Čzochralski method with a [100] axis from a charge of zone-refined germanium. The room-temperature resistivity of the charge material was greater than 40 Ω-cm which indicates that it was basically intrinsic. The dominant impurity in the charge was specified as antimony. The single crystal of silicon was produced by the floating zone process. The atmosphere was argon and the axis of the cylinder was [111]. The diameter of the "as-received" crystals varied slightly along the axis. Their crystal faces were clearly visible.

A rather elaborate machining operation was used on the crystals in order to insure that the axial heater was exactly in the center of the sample and that the thermocouples were at exactly known positions along the radii. These physical conditions imposed on the heater and thermocouples were solved by cutting the sample into two exactly equal half-cylinders. For this purpose, the original single crystal was centerless, ground to a uniform diameter and carefully cut lengthwise into two approximately equal halves. The flat side of each half-cylinder was given an optically flat polish so that the halves would fit tightly together. Thermocouple grooves were then cut in one of the halves. The nominal radial distances of the center line of the three thermocouple grooves from the axis were: 2.5 and 10 mm on one side of the heater groove, and 6.5 mm on the other side. A heater groove was cut along the axis of each half so that the heater would be in the center of the assembled cylinder. After the grooves were cut, the two halves were temporarily cemented together, but were offset slightly along the longitudinal axis. They were centerless ground a second time so that the interfacial plane and heater wire passed through the center of the cylinder.

After the machining, the sample was assembled and encapsulated in a molybdenum case with a high-purity Al_2O_3 ceramic wafer placed at each end. See Fig. 1. Molybdenum was used because it does not form any low melting alloys with either Si or Ge. Neither the Mo nor the Al_2O_3 contributes electrically active impurities to the Si or Ge in significant amounts during the lengthy, high-temperature runs.

The intrinsic electronic K could be suppressed if, inadvertently, the material became highly doped to a level of about 10^{18} carriers/cc. As a check on the impurity concentration, the electrical conductivity and Hall coefficient of the germanium and silicon samples were measured before and after the high-temperature measurements. Before use the room-temperature resistivity of the germanium crystal was 46.6 Ω-cm, and the

[2] W. C. Dash, J. Appl. Phys. 29, 736 (1958).
[3] W. C. Dash, J. Appl. Phys. 30, 459 (1959).
[4] W. C. Dash, J. Appl. Phys. 31, 736 (1960).
[5] W. C. Dash (private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

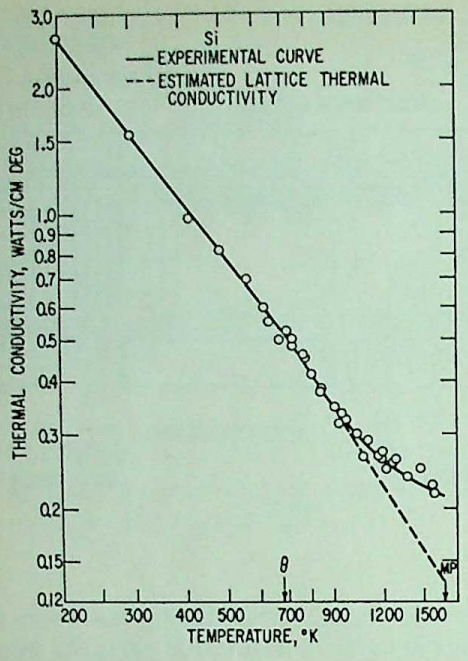


FIG. 2. K versus T for Si showing the experimental points and the extrapolated lattice component K_g. The Debye and melting temperatures are θ and MP, respectively.

net carrier concentration was 2×10^{12}/cm^3. The crystal was so lightly doped that it could not be well characterized as either n- or p-type. After the K had been measured to near the melting point, the electrical resistivity dropped to 4.6 Ω-cm and the crystal was now doped with 8×10^{15} acceptors/cm^3. These acceptors were identified as copper. The copper was probably introduced during the machining of the germanium sample prior to the measurements.

The room-temperature resistivity of the high-temperature silicon sample was 440 Ω-cm with a carrier concentration of 1.27×10^{13}/cm^3 and it was n type. After the K had been measured to near the melting point of silicon, the room-temperature resistivity dropped to 177 Ω-cm, the carrier concentration rose to 2.46×10^{13}, and it was still n type. The specific impurities could not be identified. The important conclusion from these measurements is that neither the germanium nor silicon contained enough electrically active impurities to affect the intrinsic polar or bipolar thermal conductivity in any way.

APPARATUS

The low-temperature measurements were made with a steady-state, longitudinal heat flow apparatus designed by Slack.[1] The final form of the high-temperature, cylindrical-flow thermal conductivity apparatus was obtained after several modifications. Figure 1 shows the details of the sample in the final version of the high-temperature apparatus. A more detailed description of the entire method and apparatus will be published elsewhere later. The central section of the apparatus is shown in Fig. 1. The vacuum plumbing was designed so that the sample chamber could be evacuated to 10^{-6} Torr and then filled with ultra-pure helium. The outer ceramic tube serves as the wall of the vacuum chamber for the sample. The external, 3-section electric heater is wound on this tube, and serves to control the average temperature of the sample. The split inner tube holds the sample and all of the removable parts, i.e., thermocouples, heaters, shields, etc., in place.

The cylindrical geometry for measuring the high-temperature thermal conductivity was chosen for reasons previously stated.[6] In the present, more thorough measurements, certain experimental problems arose which will now be given a cursory examination. Ultra-pure helium gas was used for the atmosphere, because it contained a low oxygen content,[7] and thus there were no oxidation problems at high temperatures. The gas served to provide good thermal contact between the sample and the heater and thermocouple wires. The gas also partially helped to suppress the vapor transport of Si and Ge, which would be appreciable at high temperatures in a vacuum. The molybdenum jacket which encased the germanium and silicon reduced the exposed surface area of the samples and hence further reduced the evaporation. This case also held the halves together. Molybdenum was employed for the jacket, axial heater wire, voltage leads, and heater leads because of its previously mentioned chemical stability. The temperature gradients were measured with platinum versus platinum-10% rhodium because these thermocouples are very stable with time and possess well-

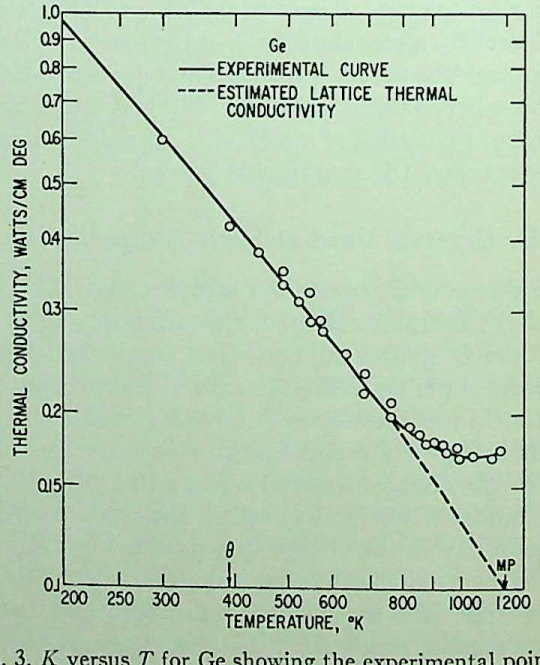


FIG. 3. K versus T for Ge showing the experimental points and the extrapolated lattice component K_g. The Debye and melting temperatures are θ and MP, respectively.

[6] G. A. Slack and C. J. Glassbrenner, Phys. Rev. **120**, 782 (1960).

[7] J. R. Young and N. R. Whetten, Rev. Sci. Instr. **32**, 453 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

known calibration curves. However, they have a very small emf output and are subject to contamination problems unless carefully used. The thermocouples and axial crystal heater were insulated from the sample by Degussit[8] alumina tubing because it is a very pure ceramic, and is available in very small diameters. It does contain a small amount of iron and silicon, i.e., about 300 ppm. The radiation shields were made from platinum −40% rhodium. These served to keep heat losses from the end of the sample low. They also served as vapor barriers for germanium and silicon vapor coming from the sample. The external heater winding was made in three sections so that both the temperature gradient along the axis of the apparatus near the ends of the sample and the average temperature of the sample could both be controlled. A ground sheath was placed under the external heater to reduce stray electrical leakage currents. The ceramic end plugs were installed to hold the electrical wires in place and to reduce axial heat conduction by the gas. They also helped to protect the thermocouple wires in the temperature gradient from the silicon and germanium vapors.

The electrical heater power to the sample was furnished by either dc or ac. Direct current was used below about 750°K and ac was used (about 435 cycles/sec) at higher temperatures. Due to the lower electrical resistance of the Si, the Ge, and the Al_2O_3 at high temperatures, a small leakage voltage from the crystal heater always appeared in the thermocouples above about 850°K. This leakage was rendered harmless by use of ac since any small 435-cycle ac voltage was not recorded by the thermocouple galvanometer. To insure that the temperature gradient was measured accurately, a few special measuring techniques were employed. Since no two thermocouples are exactly alike, two of the three thermocouples inside the sample were always calibrated in terms of the third when the sample was in an isothermal condition. Checks on the presence of possible thermocouple contamination were made in two ways. First the high-temperature K data points were always followed by a reconfirmation of lower temperature K points at about 400°K. Secondly, variations in the axial temperature gradients in the apparatus at and beyond the radiation shields were intentionally produced by the end windings on the external heater. These variations should not affect the measured K. Any apparent effect meant contaminated thermocouples. Discrepancies of these types resulted in rejection of the high-temperature data, and a reassembly of the entire apparatus with new thermocouples. The final results for both Si and Ge each involved several such reassemblies.

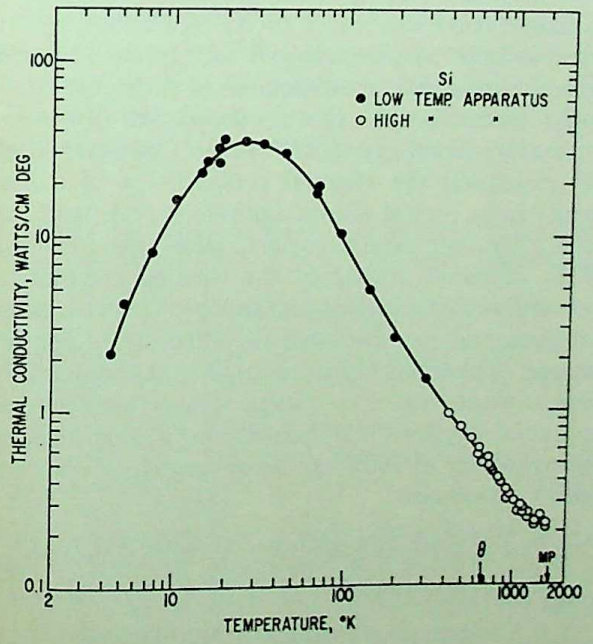


FIG. 4. The low- and high-temperature K versus T results for Si.

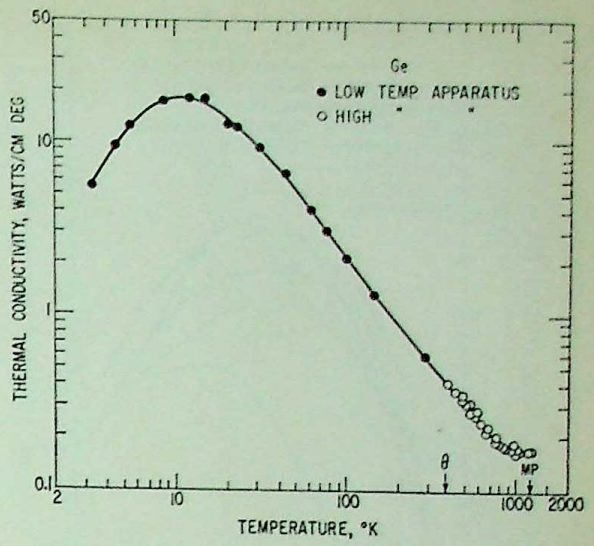


FIG. 5. The low- and high-temperature K versus T results for Ge.

Under steady-state operating conditions between 30 and 50 W of electrical power was furnished to the crystal heater. This is equivalent to about 4 W/cm of sample length. The average temperature differences between the center and surface of the sample ran between 1 and 3°K. The heater wire temperature usually operated between 50 and 100°K above the sample temperature.

RESULTS

The results of the present investigation of the high-temperature K of Si and Ge are shown in Figs. 2 and 3. The composite results for single crystal Si are shown in Fig. 4. Notice that the low-temperature results (solid circles) join on smoothly to the high-temperature results (open circles). Both sets of data were obtained in their respective apparatuses as absolute thermal conductivity values, and neither set of data has been adjusted to match the other. The actual data points in the low-temperature set span the temperature range from 4.3 to 304°K, while the high-temperature points cover the

[8] Degussit A123, Liasson Company, New York, New York.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

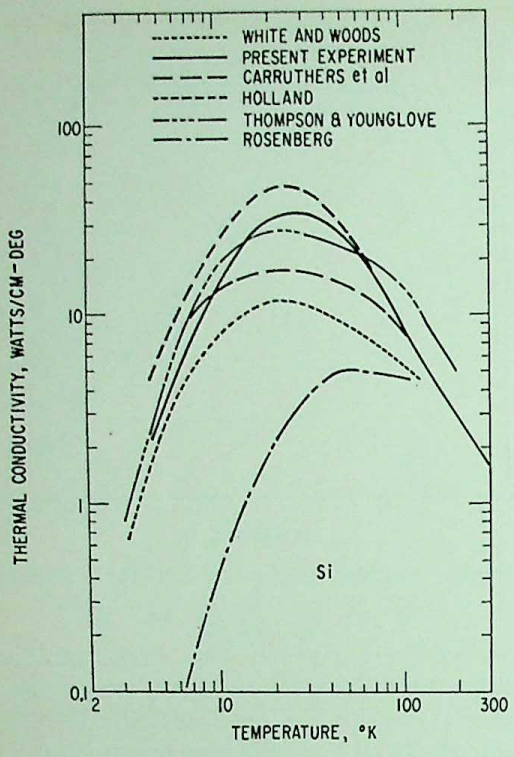


FIG. 6. Comparison of present and previous low-temperature K versus T results for Si.

418 to 1577°K range. The Debye temperature of Si at 674°K is shown for scaling purposes. The melting point, MP, of silicon at 1681°K is only 104°K above the highest point measured.

The results for single crystal Ge are shown in Fig. 5. The points for temperatures between 3.2 and 300°K have been taken from Slack and Glassbrenner.[6] The high-temperature data points between 398 and 1194°K were measured in the present high-temperature apparatus. Neither set of data was adjusted, and again the agreement between the low-temperature and high-temperature data is very good. The Debye temperature θ of Ge is 395°K. These results approach to within 16°K of the melting point, MP, of Ge at 1210°K. The absolute accuracy of both the high- and low-temperature results is computed to be ±5%, while the relative accuracy of any two points measured on the same sample is ±2%. Table I gives smoothed values for the K of Si and Ge. The low-temperature region where K depends on sample size has been excluded.

TABLE I. Representative values of the thermal conductivity K for high-purity single crystals of silicon and germanium (accuracy ±5%).

	K, W/cm deg			K, W/cm deg	
T°K	Si	Ge	T°K	Si	Ge
20		15.	400	1.05	0.44
30		10.5	500	0.80	0.338
40		7.7	600	0.64	0.269
50	26.	5.9	700	0.52	0.219
60	21.	4.7	800	0.43	0.193
70	17.0	3.7	900	0.356	0.177
80	13.9	3.1	1000	0.310	0.171
90	11.4	2.55[a]	1100	0.280	0.169
100	9.5	2.25	1200	0.261	0.173
125	6.0	1.66	1210		0.173
150	4.20[a]	1.30	1300	0.248	
175	3.25	1.10	1400	0.237	
200	2.66	0.95	1500	0.227	
250	1.95	0.73	1600	0.219	
300	1.56	0.60	1681	0.216	

[a] The bar over the last digit indicates that it is rather uncertain.

PREVIOUS DATA

A. Silicon Below 300°K

A comparison of the present and previous low-temperature K measurements of silicon below 300°K is shown in Fig. 6. Except for the measurements by Thompson and Younglove, all of the measurements agree quite well above 100°K. Some of the variation in K below 100°K is accounted for by boundary scattering in different size samples. However most of the variation in Fig. 6 is caused by phonon scattering from impurities and other crystalline defects. The earliest report of thermal conductivity measurements of silicon at low temperatures of which the authors are aware was made by Rosenberg.[9] He measured a polycrystalline silicon sample over the temperature range 2 to 100°K, and employed gas thermometers to measure the temperature gradient during steady-state heat flow. Scattering by the grain boundaries in his polycrystalline sample is probably the reason that Rosenberg's K results are much lower than any of the others. White and Woods[10] have measured a single crystal silicon sample (room-temperature resistivity 6.7 Ω-cm) over a temperature range of 2 to 100°K. The impurities present in this sample were not stated. Thompson and Younglove,[11] in the temperature interval 2 to 200°K, have measured several samples of silicon doped with boron or arsenic. Figure 6 shows their measurement of their most highly doped silicon sample (boron doped with a room-temperature resistivity of 2000 Ω-cm). Carruthers *et al.*[12] have measured the thermal conductivity of a high-purity single crystal silicon sample, n type from 2 to 300°K. The net carrier concentration was 5×10^{14} at 300°K. Holland[13] measured the thermal conductivity of several silicon samples containing different amounts of oxygen and has discussed the effect of oxygen as a scatterer of phonons to reduce the thermal conductivity. His measurements of an n-type sample with the least amount of oxygen ($<10^{16}$ cm^{-3}) and a room-temperature resistivity of 260 Ω-cm agree very closely with the present experiment.

[9] H. M. Rosenberg, Proc. Phys. Soc. (London) **A67**, 837 (1954).
[10] G. K. White and S. B. Woods, Phys. Rev. **103**, 569 (1956).
[11] J. C. Thompson and B. A. Younglove, Phys. Chem. Solids **20**, 146–149 (1961).
[12] J. A. Carruthers, T. H. Geballe, H. M. Rosenberg, and J. M. Ziman, Proc. Roy. Soc. (London) **A238**, 502 (1957).
[13] M. G. Holland, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Academic Press Inc., New York, 1962), p. 633.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

B. Above 300°K

Various previous authors have measured the K of Si and Ge over part of the high-temperature range studied here. Consider Si first. Figure 7 shows how the present results on Si for $T>300$°K compare with some of the better previous data. The earliest K data on Si that has been reported is that of Koenigsberger and Weiss[14] in 1911 who found 0.84 W/cm deg at 290°K for an impure sample. Later authors,[13–24] who have measured K as a function of temperature for $T>300$°K, are listed in Table II. The recent values of K determined by Shanks *et al.*[23] from thermal diffusivity measurements are about 10% lower than the present ones between 300 and 900°K, and 20% higher at 1400°K. The absolute accuracy of the K values in the present experiment is ±5%, and this same limit of error seems to be present in the data of Shanks *et al.*, as judged from the scatter of the points and the size of the samples used. Thus the 10% difference is just within the combined limit of error. The almost temperature-independent K found by Shanks *et al.* for $T>1000$°K, however, is not confirmed. The K of Si continues to decrease with increasing temperature up to 1580°K, The high-temperature heat-capacity data used by Shanks *et al.* is in reasonably good (±2%) agreement with that given by Kelley,[25] so the conversion from thermal diffusivity to K should be accurate. The reason for the discrepancy in K determined by these two different methods is not as yet understood. The K results of Abeles *et al.*[21,22] from 310 to 1200°K, also determined from thermal diffusivity measurements, fall between the present curve and that of Shanks *et al.*

In reviewing the literature on the K of Ge it can be seen that the early measurements on Ge by Greico and Montgomery[26] gave $K=0.59$ W/cm deg at 298°K, in good agreement with the present value of 0.60 W/cm deg at 300°K. The temperature variation of K of Ge for temperatures near or above room temperature has been measured by many investigators. These are listed[6,21,22,26–36] in Table II. Some of these K versus T curves for Ge are shown along with the present results in Fig. 8. The agreement of the older results with the present ones is to within ±20% at 300°K, but becomes increasingly poor at the higher temperatures. The very abrupt rise in K suggested by the authors[6] in 1960 is not confirmed. The change in slope is much more gradual. The large range of values for K of Ge above 700°K found in the other curves in Fig. 8 can be attributed to experimental difficulties in measuring and controlling the thermal radiation losses. The K values deduced from thermal diffusivity measurements by Abeles *et al.*[21,22] are not subject to this source of error, and the agreement of these results with the present ones between 500 and 1070°K is very good. The 10% discrepancy between the present results and those of Abeles *et al.* at 300°K is about at the outer limits of accuracy

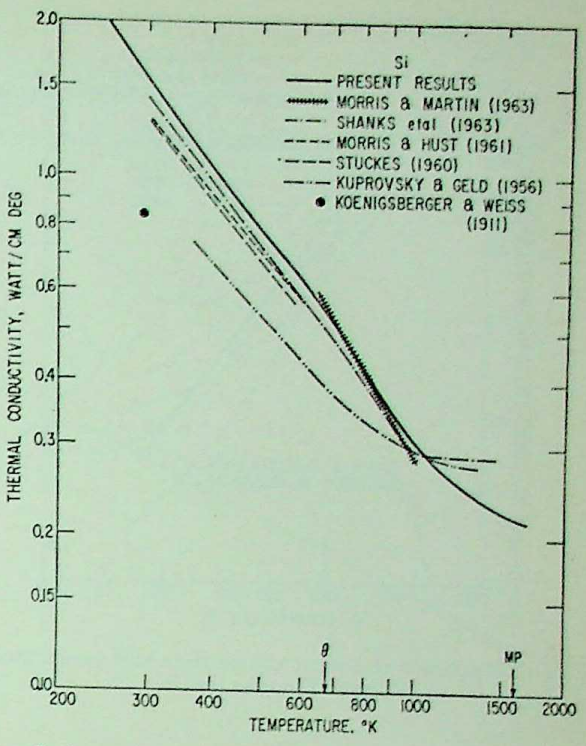


FIG. 7. Comparison of present and previous high-temperature K versus T results for Si.

---

[14] J. Koenigsberger and J. Weiss, Ann. Phys. **35**, 1 (1911).
[15] B. B. Kuprovski and P. V. Geld, Fiz. Metal i Metalloved. **3**, 182 (1956).
[16] P. V. Geld, Zh. Tekhn. Fiz. **27**, 113 (1957) [English transl.: Soviet Phys.—Tech. Phys. **2**, 95 (1957)].
[17] W. D. Kingery, J. Am. Ceram. Soc. **42**, 617 (1959).
[18] H. Mette, W. W. Gärtner, and C. Loscoe, Phys. Rev. **117**, 1491 (1960).
[19] A. D. Stuckes, Phil. Mag. **5**, 84 (1960).
[20] R. D. Morris and J. G. Hust, Phys. Rev. **124**, 1426 (1961).
[21] B. Abeles, D. S. Beers, G. D. Cody, and J. P. Dismukes, Phys. Rev. **125**, 44 (1962).
[22] D. S. Beers, G. D. Cody, and B. Abeles, *Proceedings of the International Conference of the Physics of Semiconductors* (The Institute of Physics and the Physical Society, London, 1962), p. 41.
[23] H. R. Shanks, P. D. Maycock, P. H. Sidles, and G. C. Danielson, Phys. Rev. **130**, 1743 (1963).
[24] R. G. Morris and J. L. Martin, J. Appl. Phys. **34**, 2388 (1963).
[25] K. K. Kelley, U. S. Bureau of Mines Bulletin No. 584, Washington, D. C. (unpublished).
[26] A. Greico and H. C. Montgomery, Phys. Rev. **86**, 570 (1952).
[27] K. A. McCarthy and S. S. Ballard, Phys. Rev. **99**, 1104 (1955).
[28] A. F. Ioffe, Can. J. Phys. **34**, 1342 (1956).
[29] M. Shtenbeck and P. I. Baranskii, Zh. Tekhn. Fiz. **27**, 233 (1957) [English transl.: Soviet Phys.—Tech. Phys. **2**, 205 (1957)].
[30] P. I. Baranskii and N. S. Konopliasova, Zh. Tekhn. Fiz. **28**, 1621 (1958) [English transl.: Soviet Phys.—Tech. Phys. **3**, 1493 (1958)].
[31] B. Abeles, Phys. Chem. Solids **8**, 340 (1959).
[32] F. Kettel, Phys. Chem. Solids **10**, 52 (1959).
[33] J. I. Pankove, Rev. Sci. Instr. **30**, 495 (1959).
[34] A. D. Stuckes, Phil. Mag. **5**, 84 (1960).
[35] E. D. Devyatkova and I. A. Smirnov, Fiz. Tverd. Tela **2**, 561 (1960) [English transl.: Soviet Phys.—Solid State **2**, 527 (1960)].
[36] E. D. Devyatkova and I. A. Smirnov, Fiz. Tverd. Tela **4**, 1669 (1962) [English transl.: Soviet Phys.—Solid State **4**, 1227 (1962)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

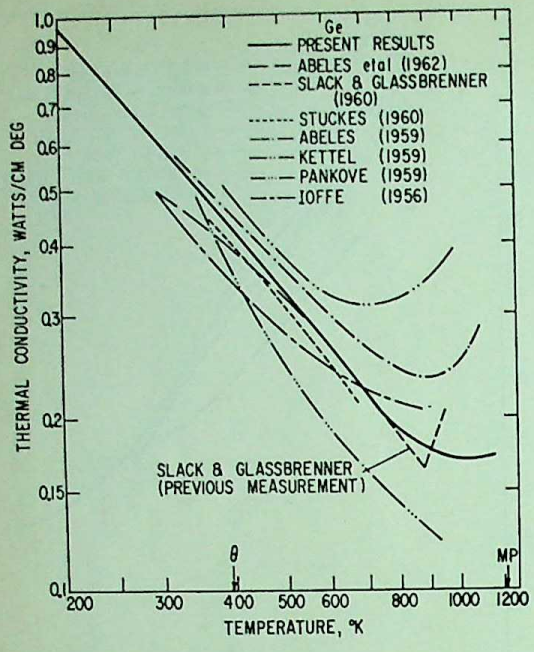


FIG. 8. Comparison of present and previous high-temperature K versus T results for Ge.

of the two experiments. The reason for this probably real discrepancy is not completely understood.

INTERPRETATION OF LOW-TEMPERATURE MEASUREMENTS

In the temperature range below 300°K, the K results on Si and Ge can be interpreted using Callaway's formalism.[37–39] The K due to heat transport by phonons is given by

$$K=\frac{k}{2\pi^2 v}\left(\frac{kT}{\hbar}\right)^3\int_0^{\theta/T}\tau_C(x,T)\frac{x^4e^x}{(e^x-1)^2}dx, \qquad (1)$$

where k = Boltzmann's constant, v = average sound velocity, T = absolute temperature, h = Planck's constant, θ = Debye temperature, $x=(\hbar\omega/kT)$, ω = phonon frequency, and τ_C is a combined relaxation time. The complicating effects[37] of normal phonon scattering have been neglected. The combined relaxation time is thus taken as

$$\tau_C^{-1}=\tau_B^{-1}+\tau_I^{-1}+\tau_U^{-1}, \qquad (2)$$

where τ_B, τ_I, τ_U are the relaxation times for boundary, isotope, and umklapp scattering of the phonons.

The particular form of the relaxation times shall now be examined. Klemens[40,41] has derived a relaxation time for scattering by point imperfections. His results are valid for isotope scattering.

$$\tau_I^{-1}(\omega)=3V_0\Gamma\omega^4/\pi v^3, \qquad (3)$$

where V_0 = average volume per atom in the crystal, and Γ = point impurity scattering parameter.[42] For umklapp scattering the following semiempirical expression[43] will be used:

$$\tau_U^{-1}=B_U\omega^2T. \qquad (4)$$

TABLE II. Literature references to the thermal conductivity of Si and Ge for $T\geq 300$°K.

Mtrl	Author	Method	Temp. range °K	Year	Ref.
Si	Koenigsberger and Weiss	c	293	1911	14
	Kuprovski and Geld	b	380–1190	1956	15, 16
	Kingery	c	370–540	1959	17
	Mette *et al.*	e	550–770	1960	18
	Stuckes	c	300–580	1960	19
	Morris and Hust	c	300–700	1961	20
	Abeles *et al.*	d	310–1200	1962	21, 22
	Shanks *et al.*	d	300–1400	1963	23
	Morris and Martin	c	680–1000	1963	24
Ge	Greico and Montgomery	c	298	1952	26
	McCarthy and Ballard	c	280–370	1955	27
	Ioffe	a	300–870	1956	28
	Shtenbeck and Baranskii	a	200–370	1957	29
	Baranskii and Konopliasova	a	80–370	1958	30
	Abeles	a	320–1070	1959	31
	Kettel	a	390–1000	1959	32
	Pankove	e	370–970	1959	33
	Stuckes	c	310–680	1960	34
	Devyatkova and Smirnov	a	80–440	1960, 62	35, 36
	Slack and Glassbrenner	a, b	3–1020	1960	6
	Abeles *et al.*	d	310–1070	1962	21, 22

[a] Absolute, steady-state, longitudinal heat flow.
[b] Absolute, steady-state, radial heat flow.
[c] Comparative, steady-state longitudinal heat flow.
[d] Variable state, thermal diffusivity.
[e] Other.

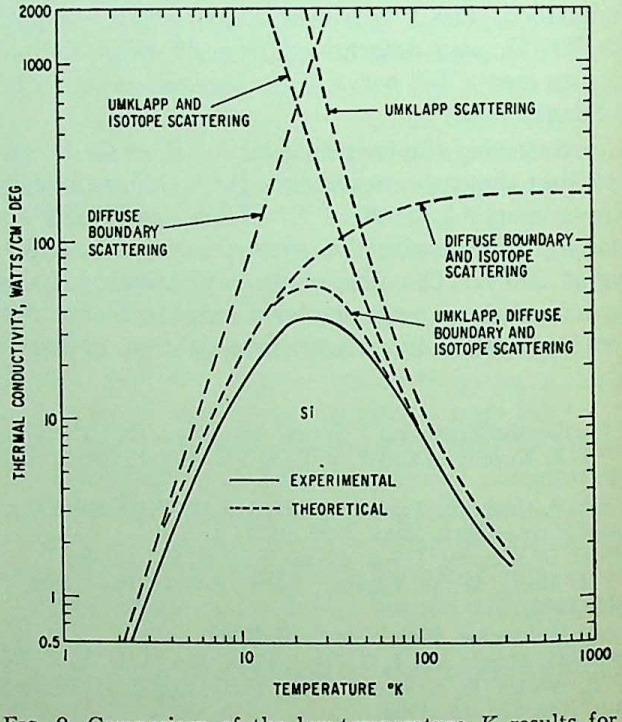


FIG. 9. Comparison of the low-temperature K results for Si with the theory in which various phonon scattering mechanisms are considered.

[37] J. Callaway, Phys. Rev. 113, 1046 (1959).
[38] J. Callaway and H. C. von Baeyer, Phys. Rev. 120, 1149 (1960).
[39] J. Callaway, Phys. Rev. 122, 787 (1961).
[40] P. G. Klemens, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1958), Vol. 7, p. 1.
[41] P. G. Klemens, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1956), Vol. 14, p. 198.
[42] G. A. Slack, Phys. Rev. 126, 427 (1962).
[43] G. A. Slack and S. Galginaitis, Phys. Rev. 133, A253 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

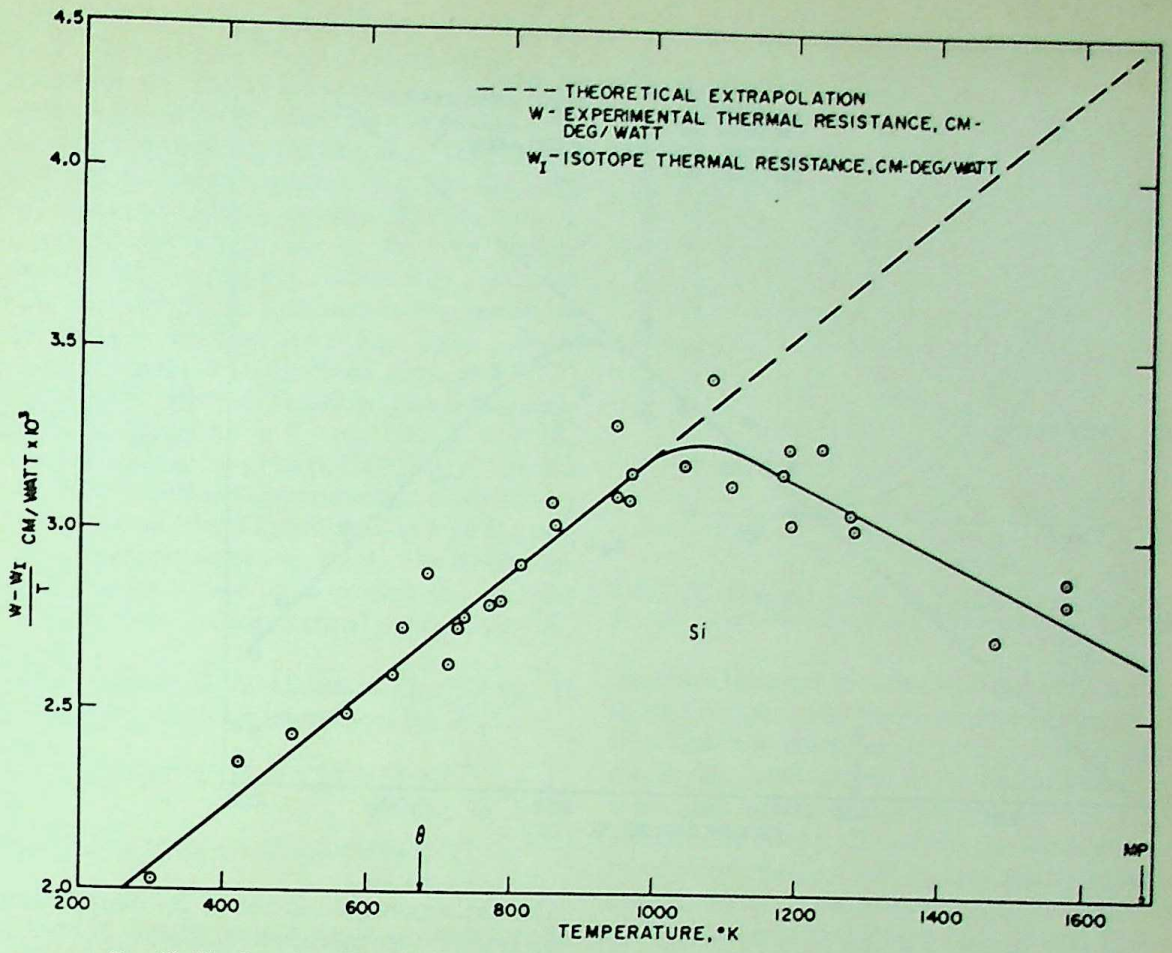


FIG. 10. The function $(W-W_I)T^{-1}$ versus T for Si as a means of determining the extrapolated lattice thermal resistivity at high temperatures.

Here B_U has the approximate value[43] of

$$B_U \sim (h\gamma^2/2\pi m\theta v^2) \exp(-\theta/3T), \quad (5)$$

where γ = Grüneisen's constant assumed to be equal to 2, and m = average mass of a single atom. For boundary scattering the relaxation time for diffuse scattering of the phonons off the walls of the crystal is simply

$$\tau_B^{-1} = v/L, \quad (6)$$

where L is the average diameter of the crystal. For Si these various relaxation times have been included in the K expression given in Eq. (1) using the following values: $v = 6.4\times10^5$ cm/sec, $\theta = 674$°K, $V_0 = 1.99\times10^{-23}$ cm^3, $\Gamma = 1.65\times10^{-5}$, and $L = 0.44$ cm. The results of the calculations for various combinations of relaxation times are shown in Fig. 9 for Si. It can be seen that all three relaxation times must be used in order to obtain agreement between theory and experiment. The theoretical curve with all three scattering processes included gives a maximum K of 56 W/cm deg at 28°K for Si compared to the observed maximum of 39 W/cm deg at 25°K. The fit at lower K values is rather better. In view of the simple model used for deriving Eq. (1) such agreement is considered reasonable. A similar sort of analysis has been carried out by other authors[6,37] for Ge at low temperatures, where the agreement is even better. Some improvement in the fitting of the theory to experiment can be made by including the effects of both the transverse and longitudinal phonon branches and the effects of phonon dispersion.[44] These added complications will not be considered here. The primary conclusion to be drawn from the low-temperature results for $T<300$°K is that phonons are the only significant carriers of the thermal energy, and they are scattered by other phonons, isotopes, and the crystal boundaries. Other carriers of thermal energy do, however, become important for $T>300$°K.

INTERPRETATION OF HIGH-TEMPERATURE MEASUREMENTS

Lattice Thermal Conductivity

In the present context high temperature means $T>300$°K. In this region it is necessary to consider photons and electrons and holes as well as phonons as possible carriers of thermal energy. We shall assume that these various carriers interact only weakly so that their contributions to the thermal conductivity are additive.

[44] M. G. Holland, Phys. Rev. 132, 2461 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

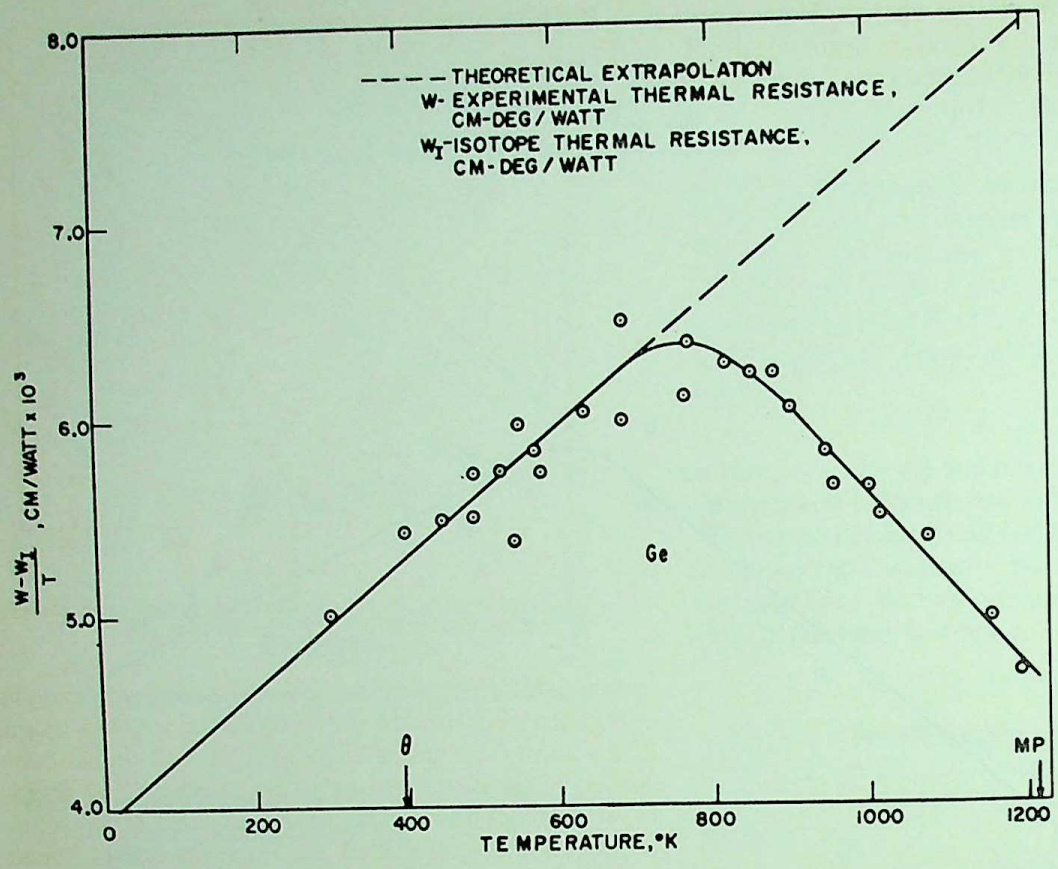


FIG. 11. The function $(W-W_I)T^{-1}$ versus T for Ge as a means of determining the extrapolated lattice thermal resistivity at high temperatures.

Thus

$$K=K_g+K_e+K_r, \tag{7}$$

where K_g is the phonon, K_e is the electronic, and K_r is the radiative or photon contribution to K.

Let us first consider the lattice thermal conductivity K_g. At high temperatures we have to consider the relaxation times τ_I and τ_U given in Eqs. (3) and (4), but we can neglect τ_B in Eq. (6). Equation (4) gives the relaxation time for 3-phonon umklapp processes. However, for $300°K<T<1681°K$, we are in the range of T comparable to or greater than the Debye temperature θ for both Si and Ge. In this range, it may be necessary to consider the relaxation times for four-phonon processes, as Pomeranchuck[45–47] has pointed out. He gives a relaxation time for these higher order (H) processes as:

$$\tau_H^{-1}=B_H\omega^2T^2, \tag{8}$$

with B_H a constant. K_g can be evaluated for $T>\theta$ from Eqs. (1), (3), (4), and (8) if τ_C^{-1} is taken as

$$\tau_C^{-1}=\tau_U^{-1}+\tau_H^{-1}+\tau_I^{-1}.$$

In the region where $T>\theta$ the quantity x in Eq. (1) is small, and the integral simplifies to

$$K=\frac{k}{2\pi^2v}\left(\frac{kT}{\hbar}\right)^3\int_0^{\theta/T}\tau_C x^2dx. \tag{9}$$

Also the exponential factor in B_U disappears to make B_U temperature-independent. Thus

$$\tau_C^{-1}=(B_UT+B_HT^2)\omega^2+A\omega^4. \tag{10}$$

For $T\geq\theta$ the isotope scattering is much less important than the phonon-phonon scattering. In this limit Eqs. (9) and (10) can be reduced by the method used by Ambegaokar,[48] to

$$K_g^{-1}\equiv W_g=W_U+W_H+W_I,$$

where

$$\begin{aligned} W_U&=\pi vhB_UT/\theta k^2,\\ W_H&=\pi vhB_HT^2/\theta k^2,\\ W_I&=4\pi^2V_0\theta\Gamma/hv^2. \end{aligned} \tag{11}$$

This reduction requires $W_g\gg W_I$. This condition is fulfilled for Si and Ge at high temperatures. The only really unknown quantity in Eq. (11) is B_H. The quantities B_U and B_H can be evaluated experimentally from a plot of $(W_g-W_I)T^{-1}$ versus T. The quantity W_I is, except for a difference of a factor of 12 in the definition of Γ, the same as that given by Ambegaokar.[48] For Si one obtains $W_I=0.033$ cm deg/W. For Ge the value is $W_I=0.17$ cm deg/W using $V_0=2.26\times10^{-23}$ cm^3, $\theta=395°K$,[49] $\Gamma=4.90\times10^{-5}$,[50] and $v=3.94\times10^5$ cm/sec.

[45] I. Pomeranchuk, Phys. Rev. **60**, 820 (1941).
[46] I. Pomeranchuk, J. Phys. USSR **4**, 259 (1941).
[47] I. Pomeranchuk, J. Phys. USSR **7**, 197 (1943).
[48] V. Ambegaokar, Phys. Rev. **114**, 488 (1959).
[49] P. Flubacher, A. J. Leadbetter, and J. A. Morrison, Phil. Mag. **4**, 273 (1959).
[50] D. Strominger, J. M. Hollander, and G. T. Seaborg, Rev. Mod. Phys. **30**, 585 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The determination of W_g is the next problem. Consider Eq. (7). As has been shown,[6] K_r is negligible at all temperatures for Ge. The measurements of Ukhanov[51] on the optical absorption of Si and Ge at elevated temperatures were employed to calculate K_r. This calculation shows that K_r is also negligible (i.e., less than 2% of total K) for Si at all temperatures. The K_e term in Eq. (7) becomes important only at the very highest temperatures when the number of thermally generated carriers becomes very large. This term is responsible for the upturn in the K results shown in Figs. 2 and 3 above 800°K. Thus K_g is the only significant term in Eq. (7) for $T<800$°K in Si and Ge. Therefore, the total measured thermal conductivity is $K=W_g^{-1}$ for $T<800$°K. Figures 10 and 11 show the plots of $(W-W_I)T^{-1}$ versus T for Si and Ge. The linear approximation predicted by Eq. (11) is satisfactory for $T\geq 0.5\theta$ until at high temperatures the K_e term becomes important. The dashed extrapolations in Figs. 10 and 11 show what the behavior would be if only the lattice thermal conductivity K_g were present.

The linear portions of the curves in Figs. 10 and 11 yield the following empirical expressions for W_g:

$$\begin{aligned}&\text{Ge:}\quad W_g=(3.95\times10^{-3}T+3.38\times10^{-6}T^2+0.17)\ \text{cm deg/W},\\&\text{Si:}\quad W_g=(1.56\times10^{-3}T+1.65\times10^{-6}T^2+0.03)\ \text{cm deg/W}.\end{aligned}\tag{12}$$

The first term in these equations yields an experimental value of B_U which agrees with that estimated from Eq. (5) to within a factor of two for both Si and Ge. Such agreement is quite satisfactory. If the very approximate theory used in obtaining Eq. (11) is taken at face value, the results in Eq. (12) for B_H show that the ratio τ_H/τ_U is 3 for Ge and 1 for Si at $T=\theta$. This ratio becomes smaller at higher temperatures. Notice that we have chosen to interpret the observation that K_g for Si and Ge at high temperatures decreases faster than T^{-1} as a sign that 4-phonon processes are taking place. A similar conclusion has been arrived at by Stuckes[19] and by Steigmeier and Kudman[52] for various IV and III-V semiconductors. The result that the 3-phonon and 4-phonon relaxation times, τ_H and τ_U, are nearly equal at $T=\theta$ should, however, be greeted with some caution. The main purpose of the plots in Figs. 10 and 11 is to allow one to make a reasonable extrapolation of K_g in the temperature region where K_e is appreciable.

This extrapolation of K_g from Figs. 10 and 11 has been replotted as the dashed curve in Figs. 2 and 3. The difference between this curve and the experimental results is, from Eq. (7), just K_e. These experimental values of K_e for Si and Ge are plotted in Figs. 12 and 13.

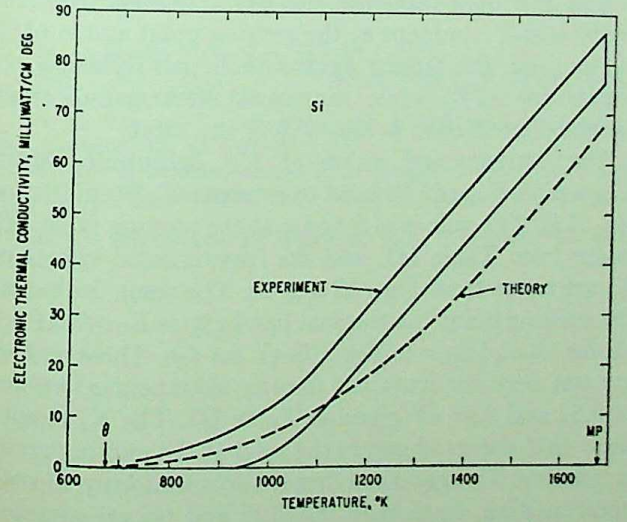


FIG. 12. Comparison of the experimental and theoretical electronic thermal conductivity, K_e versus T, for Si.

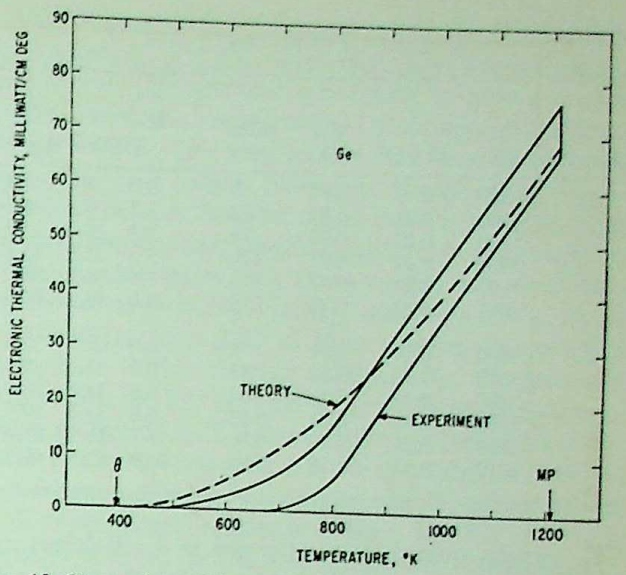


FIG. 13. Comparison of the experimental and theoretical electronic thermal conductivity, K_e versus T, for Ge.

Electronic Thermal Conductivity

The experimental results for K_e can be compared with the theory that has been developed by Davydov and Shmushkevitch,[53] by Price,[54] and by Drabble and Goldsmid.[55] Their results for K_e can be written as the sum of an electronic polar contribution K_{ep}, and an electronic bipolar part K_{eb}.

$$K_e=K_{ep}+K_{eb},$$

where

$$\begin{aligned}K_{ep}&=2\left[\frac{k}{e}\right]^2\sigma T,\\K_{eb}&=\frac{b}{(1+b)^2}\left[\frac{E_G}{kT}+4\right]^2\left[\frac{k}{e}\right]^2\sigma T.\end{aligned}\tag{13}$$

[51] Yu. I. Ukhanov, Fiz. Tverd. Tela 3, 2105 (1961) [English transl.: Soviet Phys.—Solid State 3, 1529 (1962)].

[52] E. F. Steigmeier and I. Kudman, Phys. Rev. 132, 508 (1963).

[53] B. Davydov and I. Shmushkevitch, Uspekhi Fiz. Nauk USSR 24, 21 (1940).

[54] P. J. Price, Phil. Mag. 46, 1252 (1955).

[55] J. R. Drabble and H. J. Goldsmid, *Thermal Conduction in Semiconductors* (Pergamon Press, Inc., London, 1961), pp. 104–119.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. The theoretical electronic thermal conductivity of Si and Ge versus temperature.

Crystal	T °K	$\frac{b}{(1+b)^2}$	E_G eV	σ ohm^{-1} cm^{-1}	K_{ep}	K_{eb}	K_e
					mW/cm deg		
Si	800	0.21	0.92	3.9	0.0	1.4	1.4
	1000	0.21	0.85	22.	0.3	6.2	6.5
	1200	0.21	0.78	73.	1.3	16.4	17.7
	1400	0.21	0.71	170.	3.5	31.9	35.4
	1600	0.21	0.55	315.	7.5	50.5	58.0
	1681	0.21	0.52	385.	9.6	58.2	67.8
Ge	400	0.20	0.61	0.9	0.0	0.3	0.3
	600	0.19	0.52	30.	0.3	5.0	5.3
	800	0.18	0.43	160.	1.9	18.0	19.9
	1000	0.16	0.34	440.	6.5	33.4	39.9
	1200	0.16	0.26	870.	15.5	51.9	67.4
	1210	0.16	0.25	890.	16.0	52.5	68.5

The various quantities are defined as k=Boltzmann's constant, e=electronic charge, σ=total electrical conductivity at T, b=mobility ratio at T, and E_G=indirect energy gap at T. The polar part K_{ep} is the usual Wiedemann-Franz-Lorenz contribution present in metals. The bipolar part K_{eb} is a property of semiconductors, and is caused by electron-hole pairs with energy E_G diffusing down the temperature gradient. Equation (13) has been derived for a very idealized semiconductor which is assumed to be intrinsic, to possess simple parabolic valence and conduction bands, and in which the charge carriers suffer only acoustic mode lattice scattering where the scattering probability varies with the carrier energy e as $\sqrt{e}$. Neither Si nor Ge can be said to fit this model very well since they have multiple minima conduction bands, multiple valence bands, and since both acoustic mode and intervalley scattering occur. An exact calculation of K_e in which all of these effects are considered is not undertaken here. However, as will be seen, Eq. (13) yields predictions which are in reasonable agreement with the experiment.

The experimental parameters that are used in Eq. (13) are σ, b, and E_G. Representative values for these parameters and the calculated values for K_{ep} and K_{eb} are given in Table III. For Si, the $\sigma(T)$ values have been taken from Morin and Maita[56] and Mokrovskii and Regel,[57] and the values of b versus T have been estimated from an extrapolation of the mobility curves of Morin and Maita.[56] The $E_G(T)$ values[56] have been corrected for more recent values of the effective mass, and are taken from Hannay.[58] It has been assumed that E_G varies linearly with temperature and the value of $E_G=(1.21-4.1\times10^{-4}T)$ eV was used.

For Ge the $\sigma(T)$ values have been taken from Morin and Maita,[59] from Epstein,[60] and from the recent compilation of Hamilton and Seidensticker.[61] The values of b versus T have been estimated from an extrapolation of the data given by Morin and Maita.[59] The $E_G(T)$ values[59] have been corrected for better effective mass values.[58] The equation used for Ge was $E_G=(0.785-4.4\times10^{-4}T)$ eV. The computation of K_{ep} depends only on σ and is probably accurate to $\pm5\%$. The computation of K_{eb} depends similarly on σ and is also sensitive to errors in E_G and b. The over-all accuracy in the calculation of K_{eb} is estimated to be $\pm15\%$. Figures 12 and 13 show how the calculated theoretical value of K_e compares with the experimental determination, which is judged to be accurate to within ±5 mW/cm deg. This uncertainty is indicated by the broad band in Figs. 12 and 13. Since the theoretical values of K_e are probably accurate to $\pm15\%$, the agreement with theory is considered quite adequate. In view of the uncertainties in both the calculated and experimental values of K_e it does not seem worthwhile, at present, to undertake a more exact calculation of K_e in which the details of the band structure of Si and Ge are carefully taken into account.

TABLE IV. The experimentally determined contributions to K in solid Si and Ge at their respective melting points.

Quantity	Si 1681°K	Ge 1210°K
K_{ep}/K	5%	9%
K_{eb}/K	32%	32%
K_g/K	63%	59%
K (W/cm deg)	0.22	0.17

The three contributions to the total thermal conductivity, K_{tot}, at the melting point are K_g, K_{ep}, and K_{eb}. Table IV shows the relative contributions of these three terms for Si and Ge. K_g is estimated from the extrapolated dashed curve in Figs. 2 and 3, K_{ep} is calculated form Eq. (13), and K_{eb} is taken to be equal to $K_{tot}-K_g-K_{ep}$. For both materials the lattice thermal conductivity is still dominant at the melting point and is 61% of the total, the bipolar electron-hole pair diffusion accounts for 32%, while the usual Wiedemann-Franz-Lorenz contribution is only 7% of the total.

The experimental value of K_{eb} determined from $K_{tot}-K_g-K_{ep}$ can be used to estimate E_G by means of Eq. (13). The values of b and σ at the melting point are taken from Table III, and the experimental values of K_e are taken from Figs. 12 and 13. The result is that at the melting point the thermal band gap is $E_G=0.6\pm0.1$ eV for Si and $E_G=0.26\pm0.08$ eV for Ge. These values are not very far from the linearly extrapolated values of 0.52 and 0.25 eV given in Table III. The K_e results show that the band gap in the solid at the melting point is nonzero. The gap thus drops discontinuously to zero upon melting, since both liquid Si and Ge are metallic in character.[57,61]

[56] F. J. Morin and J. P. Maita, Phys. Rev. 96, 28 (1954).
[57] N. P. Mokrovskii and A. R. Regel, Zh. Tekhn. Fiz. 23, 779 (1953).
[58] N. B. Hannay, *Semiconductors* (Reinhold Publishing Corporation, New York, 1959), p. 332.
[59] F. J. Morin and J. P. Maita, Phys. Rev. 94, 1525 (1954).
[60] E. A. Epstein, Ph.D. thesis, Purdue University, 1954 (unpublished).
[61] D. R. Hamilton and R. G. Seidensticker, J. Appl. Phys. 34, 2697 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

MAGNETIC FIELD EFFECTS

The electronic contribution to the K of Ge is about 40% of the total at 1175°K, a temperature 35°K below the melting point. An attempt was made to alter this electronic contribution by applying a moderate magnetic induction in a direction perpendicular to the axis of the K sample. Magnetic induction strengths (MKS units) up to 0.6 Wb/m² (6000 G) were applied, but no effect on K was observed to within ±2%. Steele[62] found that values up to 1.15 Wb/m² had no effect on the K of Ge at 278°K. Here the estimated $K_e=K_{ep}+K_{eb}$ is only 4×10^{-4}% of the total K for intrinsic Ge. His experiment was not sufficiently sensitive to see changes this small even if the magnetic induction had been sufficiently strong to suppress K_e. Magnetic effects on K have been seen in HgSe,[63] Bi_2Te_3,[64] and InSb.[65,66] The requirement on the magnetic induction is that it must be large enough so that the charge carriers will make about one cyclotron orbit before being scattered. This minimum value,[65] B_m, is approximately

$$B_m=\pi/\mu,$$

where B is in Wb/m², and μ is the carrier mobility in m²/V-sec. The electron mobility in intrinsic Ge at 1175°K is about 0.03 m²/V-sec, the hole mobility is even lower. The minimum magnetic induction is thus about 100 Wb/m² (10^6 G). It is rather clear why no magnetic effects were seen in Ge. No experiments on Si were tried.

CONCLUSIONS

The thermal conductivity K of silicon has been measured from 3 to 300°K using a steady-state, longitudinal heat flow apparatus. By using the relaxation times for umklapp scattering, isotope scattering and boundary scattering, and Callaway's simple model, a theoretical phonon thermal conductivity was calculated and was found to agree with the experimental measurements.

A high-temperature thermal conductivity apparatus employing steady-state, radial heat flow in a cylindrical sample was designed and was found to give accurate (±5%), absolute, measurements at temperatures from 300 to 1600°K. The method does not require any correction for heat loss by radiation. The K values so obtained agree with those found for both Si and Ge using the low-temperature, longitudinal heat flow apparatus.

The results show that phonons are the dominant carriers of heat in high purity, solid Si and Ge at all temperatures from 3°K to their melting points. The phonon or lattice thermal conductivity K_g decreases faster than the T^{-1} law predicted by three phonon umklapp scattering for $T\geq\theta$, where θ is the Debye temperature. For temperatures $T>\theta$, an approximate theoretical formula has been developed for K_g which agrees with experimental results between $0.5\theta<T<1.6\theta$. This formula is based on the assumption that three-phonon, four-phonon, and isotope scattering are important at high temperatures, and it predicts that the thermal resistivities of these three processes are additive. It is suggested that the relaxation times for both three- and four-phonon processes are comparable for $T\geq\theta$.

For temperatures above about 1.6θ an electronic thermal conductivity K_e has been found in both Si and Ge. The theoretical and experimental values of K_e agree quite well. The major part of K_e comes from the bipolar term, with some small contribution from the polar term. At the melting point the phonons contribute an average of 61% to the total K while the bipolar and polar contributions are 32% and 7%, respectively.

The thermal band gaps at the melting points are estimated from K_e as 0.6±0.1 eV for Si and 0.26±0.08 eV for Ge. They are in reasonable agreement with the linear extrapolations from lower temperatures.

ACKNOWLEDGMENTS

The authors wish to thank C. A. Reynolds and F. H. Horn for their interest in this work, and D. Hartman for his aid in obtaining the large single crystals of Si and Ge. The assistance of J. H. McTaggart in many phases of the measurements is greatly appreciated. The authors also thank M. G. Holland for a preprint of his paper on silicon and germanium, and G. C. Danielson and R. W. Powell for several interesting discussions concerning thermal conductivity measurements.

[62] M. C. Steele, Phys. Rev. 107, 81 (1957).
[63] Kh. I. Amirkhanov, A. Z. Daibov, and V. P. Zhuze, Dokl. Akad. Nauk SSSR 98, 557 (1954).
[64] A. E. Bowley, R. Delves, and H. J. Goldsmid, Proc. Phys. Soc. (London) 72, 401 (1958).
[65] D. Kh. Amirkhanova and R. I. Bashirov, Fiz. Tverd. Tela 2, 1597 (1960) [English transl.: Soviet Phys.—Solid State 2, 1447 (1961)].
[66] R. G. Morris and F. Hornstra, Rev. Sci. Instr. 33, 1067 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Soft X-Ray Production by 1.5-MeV Protons

W. T. Ogier, G. J. Lucas,* J. S. Murray,* and T. E. Holzer*
Pomona College, Claremont, California
(Received 31 December 1963)

Thick-target yields of x rays with wavelengths from 1.5 to 44 Å have been measured as functions of bombarding proton energy near 1.5 MeV. X-ray production cross sections computed from these yields are in satisfactory agreement with existing theoretical excitation cross-section calculations insofar as current estimates of fluorescence yields permit comparisons to be made. Yields at 1.5 MeV of the ultrasoft x rays measured were, in quanta per proton steradian: Al_K, 3.5×10^{-3}; Cu_L, 9×10^{-4}; C_K, 4.5×10^{-3}.

INTRODUCTION

INTEREST in determination of x-ray yields from proton-bombarded targets revived with the advent of scintillation counting. Extensive early investigations by Lewis, Simmons, and Merzbacher[1] and by Bernstein and Lewis[2] were followed by a number of experimental and theoretical studies. The state of the field was then reviewed and analyzed in a satisfying way for proton energies above one MeV by Merzbacher and Lewis.[3] Subsequent interest has centered on energies below one MeV for which the Born approximation does not yield completely satisfactory results.[4,5] The use of scintillation counters in these modern measurements served to limit observations to x rays with quantum energies of 4.5 keV or more.

The availability of thin-window proportional counter techniques makes possible quantitative study of yields of x rays with quantum energies at least as low as 280 eV.[6] Reports of such proton-yield measurements have only recently begun to appear.[7] In this investigation yields of a number of K and L x-ray bands with quantum energies between 8000 and 280 eV (1.5 to 44.6 Å) were measured for bombarding proton energies between 1.2 and 1.6 MeV. These yield functions are used to calculate cross sections for x-ray production at 1.5 MeV. The ratio of production cross section to excitation cross section is the fluorescence yield. Current estimates of excitation cross section and fluorescence yield both involve uncertainties in the soft x-ray region. These results do not resolve either uncertainty, therefore, but they fit well with reasonable estimates of both parameters.

EXPERIMENTAL PROCEDURE

Targets of the metals investigated were bombarded inside the main tank of the 23-in. cyclotron. This cyclotron, a prototype model, employs radial sector Thomas focusing without the usual tapering of the gap at larger radii to maintain orbit synchronism. Since the sector focusing is strong in the vertical direction for eccentric orbits as well as for symmetric orbits, a target inside the tank is exposed to a considerable spectrum of energies which can be limited somewhat by careful collimating of the incident beam at the target position. Such collimation was provided, and since negligible currents are found in the symmetric orbits at large radii, the angle of collimation was set at approximately 83° with the radial direction in order to select the more energetic eccentric orbits. The proton energy was varied by moving the target assembly radially. In the absence of computed energies, the yield of tantalum L x rays was used to calibrate the beam energy using the data of Bernstein and Lewis[2] extrapolated below 1.5 MeV with the aid of the calculations of Merzbacher and Lewis.[3] The surface of the target was set optically in each case at the angle of specular reflection between the beam direction and the radial direction along which x rays were viewed. The proton and the resulting x ray thus traveled equal distances in the target material.

The x rays were counted outside the magnetic field at the end of a 129-cm evacuated drift path by means of a flow proportional counter of standard design using P-10 gas (10% methane, 90% argon).[8] The window aperture was a single 0.063-in.-diam hole inside the vacuum, which thus subtended a solid angle at the target of 1.21×10^{-6} sr. The window material, 0.00025-in. aluminized Mylar, was bonded with epoxy cement over a slightly larger aperture set just behind the window. One corner of the Mylar was folded under to ground the aluminized surface. This structure gave good counting rates with proton currents smaller than one microampere in all cases except that of tantalum L radiation. Because of the importance of the tantalum data for calibration, it was checked using an aluminum window of larger aperture. The two runs were in very close agreement.

---

* National Science Foundation undergraduate research participant.

[1] H. W. Lewis, B. E. Simmons, and E. Merzbacher, Phys. Rev. 91, 943 (1953).

[2] E. M. Bernstein and H. W. Lewis, Phys. Rev. 95, 83 (1954).

[3] E. Merzbacher and H. W. Lewis, *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. 34, p. 166.

[4] S. Messelt, Nucl. Phys. 5, 435 (1958).

[5] R. C. Jopson, H. Mark, C. D. Swift, and M. A. Williamson, Phys. Rev. 131, 1165 (1963); 127, 1612 (1962) contains a recent bibliography.

[6] R. M. Dolby, Brit. J. Appl. Phys. 11, 64 (1960).

[7] J. M. Khan and D. L. Potter, Phys. Rev. 133, A890 (1964).

[8] C. F. Hendee, S. Fine, and W. B. Brown, Rev. Sci. Instr. 27, 531 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Yields were determined by counting the pulses corresponding to the x-ray band desired with a single-channel pulse-height analyzer while proton current was being monitored with a calibrated galvanometer. The pulses corresponding to a particular band were well resolved, but the structure within bands was not visible. The *bremsstrahlung* background was apparently very small and was ignored. In particular, in the case of carbon-K radiation a clear minimum existed in the pulse-height spectrum between the counting pulses and the amplifier noise. The discrimination level was set just above the noise. The yields reported from the x rays with quantum energy greater than the ionization potential of argon include the attenuated argon-K escape pulses.

Counter efficiencies due to window and gas absorption were computed from the measured geometry of the counter using the mass absorption-coefficient interpolations of the data of Allen by Henke, White, and Lundberg.[9] The path length in the argon was 2.26 cm. The computed efficiencies are recorded in the first column of Table I. The transmission factor of the Mylar for carbon-K radiation was thus estimated to be 0.04. This transmission was checked experimentally by placing a second layer of the same material in the x-ray path in the vacuum. The measured transmission varied from place to place in the film. The value adopted was 0.038 ± 0.01.

EXPERIMENTAL RESULTS

The yields of x rays, Y_μ, in quanta per proton steradian at 1.5-MeV bombarding energy are given in the second column of Table I and the yields as functions of bombarding energy are presented graphically in Fig. 1. Considerable uncertainty lies in the determination of proton energy, since the collimated proton beam had a spread in energy of some 0.2 MeV and the energy calibration below 1.5 MeV is based on extrapolation. The curves for tantalum-L and iron-K radiation both fit well the results of Merzbacher and Lewis,[3] but there is a sizable discrepancy in the data for titanium K. In the preliminary phases of this work the yields from various titanium targets were compared with those from tantalum and iron, always with essentially the results reported.

TABLE I. Data at 1.5 MeV.

	Counter efficiency	Y_μ quanta / proton·sr	μ/ρ	σ_x (1.5 MeV) cm²	Yield factor
C_K	0.038	4.5×10^{-3}	3000	3×10^{-21}	0.007
Al_K	0.45	3.5×10^{-3}	330	1.0×10^{-21}	0.065
Ti_K	0.84	3.2×10^{-4}	114	1.1×10^{-22}	0.24
Fe_K	0.54	1.0×10^{-4}	71.2	3.8×10^{-23}	0.35
Cu_K	0.31	4.5×10^{-5}	50.9	2×10^{-23}	0.4
Cu_L	0.057	9×10^{-4}	1000	1.3×10^{-21}	0.01
Mo_L	0.55	4×10^{-4}	630	6×10^{-22}	0.06
Ag_L	0.40	3.2×10^{-4}	460	4.0×10^{-22}	0.10
Ta_L	0.31	2.5×10^{-5}	150	2.7×10^{-23}	0.30

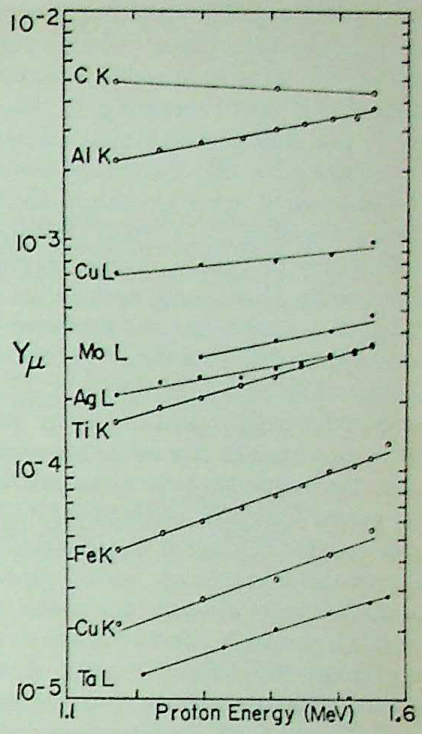


FIG. 1. Thick-target characteristic x-ray yields, Y_μ, in quanta per proton sr as functions of proton bombarding energy. Proton and quantum traveled equal distances in the target material.

The cross section for x-ray production $\sigma_x(E)$ is computed from the yield curve by the familiar formula[3]

$$\sigma_x(E)=\frac{4\pi}{n}\left(\mu Y_\mu+\frac{dY_\mu}{dE}\frac{dE}{dx}\right), \qquad (1)$$

where n is the number of atoms per unit volume and μ the linear absorption coefficient. The stopping powers, dE/dx, are readily available,[10] but the self-absorption coefficients μ are not nearly so easy to obtain, particularly for the L x rays. The K self-absorption coefficients used are based on the Henke, White, and Lundberg interpolation formula[9] and the L coefficients on the Norelco extrapolation of the data of Allen and others.[11] Even the data itself is open to question.[12] The estimated values of μ/ρ used are recorded in the third column of the table, and the resulting calculated x-ray production cross sections in the fourth column.

DISCUSSION

Theoretical excitation cross sections were derived by Bethe and Walske[13] on the basis of screened hydrogenic

[9] B. L. Henke, R. White, and B. Lundberg, J. Appl. Phys. **28**, 98 (1957).

[10] W. Whaling, *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1958), Vol. 34, p. 193.

[11] S. J. M. Allen *et al.*, Norelco Reporter, May-June 1962 (unpublished).

[12] B. A. Cooke, K. A. Pounds, P. C. Russell, and E. A. Stewardson, Proc. Phys. Soc. (London) **79**, 883 (1962).

[13] M. C. Walske, Phys. Rev. **101**, 940 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

wave functions. The resulting integrals depend critically on the minimum energy which must be transferred to the atomic state in question in order to produce excitation. The effect of screening on this energy is expressed by θ, the ratio of the ionization potential of the state in question to the theoretical ionization potential in the absence of outer screening. Merzbacher and Lewis[3] have computed and plotted excitation cross section as a function of proton energy for various values of θ. Comparing their computed curves with experimental data, they found that the predicted cross sections were too large and that the data was consistent with empirical values of θ_K and θ_L of 0.85 and 0.7, respectively, both essentially independent of atomic number. Both values correspond to effective minimum energy transfers higher than the ionization potentials, particularly for the light elements. The extrapolation of this result from the 1.5-Å wavelengths at which it was found to longer wavelengths is a doubtful procedure. Some support may be found in observations of effective thresholds for electron excitation. For example, Dolby[6] observes effective bombarding energy thresholds for strong K x-ray production of about 2200 and 400 eV from aluminum and carbon, respectively. Both correspond roughly to unit values of θ_K. The ionization potential does not appear to be a valid effective minimum for rapid energy transfer.

The ratios of observed x-ray production cross sections to theoretical excitation cross sections obtained from the curves for $\theta_K=0.85$ and $\theta_L=0.7$ plotted by Merzbacher and Lewis[3] are tabulated as yield factors in the last column of the table. To the extent that confidence can be placed in the empirical values of θ these factors are fluorescence yields. Current estimates of soft K-fluorescence yields are classifiable into a low set of values based on extrapolation of measurements with heavier metal targets[14,15] and a considerably higher set of values based on data from argon[16] and a number of theoretical calculations including recently a screened hydrogenic wave function calculation by Callan.[17] The K yield factors in the table from copper through aluminum fit well slightly below Callan's curve. Very little data on L fluorescence yields is available.[18] Since the trajectory of a massive particle in a target is inherently easier to handle theoretically than that of an electron, measurements such as these are a potentially useful way of determining fluorescence yields. They would be worth repeating with more precise apparatus.

The yield of aluminum-K x rays at 1.5 MeV corresponds to a production efficiency of 1.4×10^{10} quanta per joule sr, about the maximum efficiency obtainable. On the other hand, 1.5 MeV is about six times the proton energy required for maximum production efficiency from carbon. The maximum efficiency from carbon is readily estimated with the aid of the measured yield and the excitation cross section curve to be approximately 10^{11} quanta per joule sr. The carbon efficiency is comparable to the efficiency attainable with electron bombardment,[19] but the aluminum efficiency is smaller by a factor of four or more. The proton efficiencies can be raised somewhat by irradiating with protons at grazing incidence. Since the resulting radiation is nearly free of *bremsstrahlung*, it appears that proton excited sources might be superior to filtered electron excited sources in a few applications.

ACKNOWLEDGMENTS

We are indebted to James Backman, Elliott Bloom, and William Soldwisch for valuable contributions to various phases of this work as National Science Foundation undergraduate research participants. We also acknowledge the contribution of Professor Burton L. Henke, without whose advice and encouragement this project would not have been possible.

[14] E. H. S. Burhop, J. Phys. Radium **16**, 625 (1955).
[15] H. L. Hagedoorn and A. H. Wapstra, Nucl. Phys. **15**, 146 (1960).
[16] T. Watanabe, H. W. Schnopper, and F. N. Cirillo, Phys. Rev. **127**, 2055 (1962).
[17] E. J. Callan, Phys. Rev. **124**, 793 (1961).
[18] A. A. Konstantinov, T. E. Sazonova, and V. V. Perepelkin, Bull. Acad. Sci. USSR, Phys. Ser. **24**, 1472 (1960).
[19] A. J. Campbell, Proc. Roy. Soc. (London) **A274**, 319 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Self-Compensation Limited Conductivity in Binary Semiconductors. I. Theory*

G. MANDEL
IBM Watson Research Center, Yorktown Heights, New York
(Received 23 December 1963)

The techniques of Kroger and Vink and Brebrick are extended to allow a calculation of the minimum extent of self-compensation by simple vacancies or interstitial atoms in heavily doped binary semiconductors. The resulting equations are applied to a series of compounds, and it is found that the degree of self-compensation by singly ionizable vacancies varies from essentially complete in KCl (all but $\sim 10^{-9}$ of the impurities compensated) to practically none in GaAs (only $\lesssim 10^{-3}$ of the impurities compensated). The II-VI compounds occupy an intermediate position with about ~ 99 and 99.9% self-compensation in CdTe and ZnTe, respectively. These theoretical conductivity limitations are not sufficient to account for the experimental limitations found in, for example, *n*-ZnTe or *p*-CdS. The above results are extended to include multiply ionizable vacancies, the ionization levels of which fall within the bandgap. It is found that essentially complete self-compensation by a combination of singly and doubly ionized vacancies will occur in the higher bandgap II-VI compounds. As a consequence, for example, the Fermi level in ZnTe cannot be pushed closer to the bottom of the conduction band than half the energy separation between the second ionization level of the acceptor vacancy and the bottom of the conduction band. Some specific implications of the above calculations with respect to CdTe and GaAs are discussed. Finally, certain solubility effects (of impurities) related to stoichiometry and the above calculations are discussed.

I. INTRODUCTION

IT is well known that the introduction of electrically active impurities into a semiconductor host crystal induces the formation of electrically active "natural" defects (vacancies, interstitial atoms, etc.), which tend to at least partially compensate the electrical activity of the impurity. This phenomenon has been considered by a number of workers, particlarly in a review article by Kroger and Vink[1] and, somewhat more rigorously and recently, by Brebrick.[2,3] This self-compensation may be analyzed simply in terms of an energy balance equation, i.e., energy must be supplied by the crystal to produce the "excess" concentration of defect centers while energy is gained by the crystal by the interaction of the defects with the free carriers produced by the added impurity centers. Clearly, if the energy of defect formation is large compared to the energy gained by compensation, very little self-compensation will take place. On the other hand, if the energy of defect formation is small compared to the energy gained by compensation, all free carriers will be compensated by the formation of defects and only insulating crystals will be accessible by equilibrium processes.

The purpose of this paper is the calculation, in terms of the energy parameters of the host crystals, of the minimum extent of self-compensation by simple vacancies in a series of binary compounds, i.e., KCl, ZnTe, CdTe, and GaAs. Qualitatively, we expect to observe a trend in this series due to the fact that the more ionic compounds, e.g., KCl, have electronic bandgaps which are large compared to their cohesive energies while the reverse is true for the more covalent compounds, e.g., GaAs. Roughly speaking, the energy gained by the system upon "recombination" of a free carrier at a vacancy is expected to correlate with the bandgap, while the energy required to generate a vacancy, i.e., the energy required to remove an atom from the lattice, is expected to correlate with the cohesive energy.

The possibility of obtaining appreciable *n*- and *p*-type conductivity in the II-VI compounds, e.g., CdTe and ZnTe, is a question of increasing technological importance because of possible application toward an injection luminescence device operating in the visible region of the spectrum. These compounds appear to be an intermediate case requiring somewhat more careful analysis.

In Sec. II, we calculate the degree of self-compensation by simple singly ionizable acceptor vacancies in a binary semiconductor *MN*, into which we have introduced a large concentration of shallow donors, *D*. (A precisely analogous situation will exist for acceptors.) This calculation is based essentially upon the equation for charge neutrality and Fermi statistics as applied to a nondegenerate semiconductor.[1–3]

In Sec. III, we apply the resulting equations to the series of compounds under consideration and find the expected trend in the results. Of particular interest are the results for CdTe and ZnTe, which indicate that the degree of self-compensation by singly ionizable vacancies calculated theoretically is *not* sufficient to account for the observed conductivity limitations in the II-VI compounds, ZnTe being a particularly appropriate example.

In Sec. IV, we generalize the results of Sec. II to in-

* This research was partially supported as part of Project DEFENDER under the joint sponsorship of the Advanced Research Projects Agency, the U. S. Office of Naval Research, and the Department of Defense.

[1] F. A. Kroger and H. J. Vink, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1956), Vol. III, p. 310.

[2] R. F. Brebrick, Phys. Chem. Solids 4, 190 (1958).

[3] R. F. Brebrick, Phys. Chem. Solids 18, 116 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

clude multiply ionizable vacancies, in particular, doubly ionizable vacancies, such as might be expected to be important in the II-VI compounds. It is demonstrated that multiply ionizable vacancies can be of critical importance in determining the degree of self-compensation under certain circumstances, these circumstances being most probably appropriate to the II-VI compounds. In particular, under these circumstances, we find that the Fermi level cannot be pushed appreciably closer to the bottom of the conduction band than half the energy separation between the second ionization level of a simple doubly ionizable acceptor vacancy and the bottom of the conduction band.

Some specific implications of Secs. II and IV are discussed with respect to GaAs and CdTe in Sec. V. In particular, it is predicted that the concentration of simple vacancies in GaAs must be small, and certain experimental evidence supporting this view is cited. With respect to CdTe, it is pointed out that the assignment of a level much below the conduction band to the second ionization of a Cd vacancy cannot be correct.

Finally, in Sec. VI, we discuss certain effects of vacancies on the solubility of impurities, in view of the above calculations.

II. SINGLY IONIZABLE VACANCIES

Let us consider a binary compound semiconductor, MN, into which we have introduced a large concentration of donors, D, at the temperature T. The pressures of the various species in the vapor phase are fixed, particularly the pressures of $M_{(g)}$ and $N_{(g)}$, p_M and p_N. The only defects in the host crystal which we consider are simple singly ionizable vacancies, an M atom vacancy being assumed to be an acceptor while an N atom vacancy is assumed to be a donor. For simplicity, we will analyze only the case of *heavy* doping in which the concentration of ionized donors, D^+, is large compared to the concentration of ionized N vacancies, V_N^+, or holes in the valence band, p. [The calculations can easily be extended to include V_N^+ by simply writing $(D^+ + V_N^+)$ in place of D^+ throughout.]

It will be shown in Sec. III that self-compensation by interstitial atoms rather than vacancies is essentially equivalent and leads to similar results.

The charge conservation equation, which is sufficient to describe the system[1–3] is given by

$$D^+ = n + V_M^-, \tag{1}$$

where n is the concentration of electrons in the conduction band and V_M^- is the concentration of ionized M vacancies. The competition between n and V_M^- in maintaining charge neutrality is clearly the question at hand.

The interactions between defects in equilibrium in a nondegenerate semiconductor are given by the equations of Fermi statistics

$$n = N_c \exp(E_F - E_c/kT), \tag{2}$$

$$V_M^-/V_M = g_M \exp(E_F - E_A/kT), \tag{3}$$

where N_c is the density of states near the bottom of the conduction band (given in terms of the electronic effective mass and T), E_F is the Fermi level, E_c is the energy at the bottom of the conduction band, V_M is the concentration of neutral M vacancies, E_A is the acceptor energy level associated with an M vacancy, and g_M is a degeneracy factor (generally we assume $g_M = 2$ for a singly ionizable level). An equivalent description is given by the mass action formalism.[1–3]

$$nV_M/N_c V_M^- = 1/g_M \exp(E_A - E_c/kT), \tag{4}$$

where the right-hand side of Eq. (4) is an equilibrium constant describing the ionization of an ionized M vacancy to yield a neutral M vacancy and an electron in the conduction band.

Combining Eqs. (1) and (4), we obtain immediately

$$n/D^+ = 1/1+Q, \tag{5}$$

$$Q = g_M V_M/N_c \exp(E_c - E_A)/kT = g_M V_M/N_c \times \exp(E_g - E_A')/kT, \tag{6}$$

where E_g is the electronic bandgap and E_A' is the acceptor level of the M vacancy relative to the valence band.

Equations (5) and (6) are our central result. All the quantities in Eq. (6) except V_M, the concentration of neutral M vacancies, are generally known with sufficient accuracy (with the possible exception of E_A' in some cases). We note that, except for the temperature, T, only V_M is under experimental control. V_M will be determined by the pressure of $M_{(g)}$ in the system, p_M, which can take on a range of values. The connection between V_M and p_M is written[1–3]

$$p_M V_M/N_M = \exp \Delta S_M/k \exp(-\Delta H_M/kT), \tag{7}$$

where N_M is the concentration of M sites in the lattice, and where the right-hand side of Eq. (7) is an equilibrium constant referring to the process in which an atom of M is removed from the MN lattice and placed into the vapor, the pressure of $M_{(g)}$ being maintained at 1 atm (standard state), leaving behind an M vacancy in the lattice. The quantities ΔS_M and ΔH_M are defined as the entropy and enthalpy of M vacancy formation, respectively. Note that we have not assumed that $M_{(g)}$ is the only species in the vapor phase but are simply using p_M as a convenient measure of the chemical potential of M in the system.

Inasmuch as we wish to calculate the minimum extent of self-compensation, we must estimate the minimum value of V_M and, therefore, the maximum value of p_M. Clearly p_M must always be less than the value associated with the presence of a pure condensed phase of M (generally $M_{(l)}$ in cases of interest).[3] For temperatures

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

several hundred degrees centigrade below the maximum melting point of MN, or lower, the maximum achievable value of p_M is fairly well approximated by that associated with $M_{(l)}$, $p_M{}^{\text{vap}}$, for most cases of interest. We will confine our calculations to such temperatures that $p_M \cong p_M{}^{\text{vap}}$ and simply point out that the results will represent a lower bound to the degree of self-compensation at other temperatures where $p_M < p_M{}^{\text{vap}}$, e.g., the maximum melting point. We write,

$$p_M{}^{\text{vap}} = \exp \Delta S_M{}^{\text{vap}}/k \exp(-\Delta H_M{}^{\text{vap}}/kT), \quad (8)$$

and obtain, finally, from Eqs. (6)–(8),

$$kT \ln Q^{\min} = E_g - E_A' - \Delta H_M + \Delta H_M{}^{\text{vap}} + T(\Delta S_M - \Delta S_M{}^{\text{vap}}) + kT \ln(g_M N_M/N_c). \quad (9)$$

Thus, with Eqs. (5) and (9) we can calculate the maximum relative concentration of free carriers that we can introduce into $MN_{(s)}$ at the temperature T. If T is sufficiently large, these carriers will be equal in concentration to the "uncompensated" donor impurities present. If the temperature at which we measure the conductivity is much less than T, some of these carriers may "freeze out." The extent of this "freeze-out" is determined largely by the energy level associated with the donor impurity (for the case of singly ionizable vacancies only). We will not be concerned with this "impurity limitation" on the accessable carrier concentration and conductivity but merely state that Eqs. (5) and (9) allow us to calculate the maximum relative concentration of uncompensated shallow impurities (virtually all ionized at temperature T) that can be introduced into $MN_{(s)}$.

III. CALCULATION OF RESULTS; KCl, ZnTe, CdTe, AND GaAs

We now apply Eqs. (5) and (9) to p-KCl, n-ZnTe, n- and p-CdTe, and n-GaAs. The energy bandgaps, E_g, are available in the literature.[4–9] The energy levels of the compensating vacancies, E_A' are available for p-KCl (F center),[10] n-ZnTe,[11] n-CdTe,[8] and p-CdTe.[8,12] No data are available on E_A' in GaAs and we will simply assume $E_A' = 0$ in this case to put an upper bound on the extent of self-compensation.

The enthalpy of compensating vacancy formation, ΔH_M, is available as such only for n-CdTe.[8] The enthalpy of formation of interstitial Cd atoms in p-CdTe is reported,[8] however, according to the equation

$$I_{\text{Cd}}/N_{\text{Cd}} p_{\text{Cd}} = \exp - \Delta S_{\text{Cd}}{}^I/k \exp \Delta H_{\text{Cd}}{}^I/kT, \quad (10)$$

where I_{Cd} is the concentration of interstitial Cd atoms. It is, in practice, impossible to distinguish between interstitial Cd atoms and Te vacancies by electrical measurements.[1] If we reinterpret the data in terms of Te vacancies and use Eq. (7) (with M = Te) in place of Eq. (10), it is easy to demonstrate that the "equivalent" enthalpy of Te vacancy formation is given as

$$\Delta H_{\text{Te}} = \Delta H_{\text{CdTe}} - \Delta H_{\text{Cd}}{}^I, \quad (11)$$

where ΔH_{CdTe} is the standard molar enthalpy change for the reaction

$$\text{CdTe}_{(s)} = \text{Cd}_{(g)} + \text{Te}_{(g)}. \quad (12)$$

ΔH_{CdTe} is essentially twice the cohesive energy per gram-atom of $\text{CdTe}_{(s)}$ and is available in the literature.[13,14] On the basis of such reinterpretation, we have obtained ΔH_{Te} for the calculation appropriate to p-CdTe. It is apparent that, for our purposes, no significant difference exists between interstitial atoms and vacancies and that the available data can always be interpreted without error in terms of vacancies, as above.

The enthalpy of vacancy (F center) formation in KCl has not been reported. Data are available, however, as to the concentration of F centers as a function of potassium pressure, p_{K}, at one temperature (697°C).[15] In particular, the concentration of F centers is reported to be proportional to $p_{\text{K}}{}^{0.85}$, rather than the expected linear dependence. We may handle these data in several ways, with slight variation in the calculated results. Firstly, we arrive at the minimum value of V_{Cl} by extrapolating the reported data to a value of p_{K} consistent with the highest possible value of p_{Cl} ($p_{\text{Cl}}{}^{\text{vap}}$), an extrapolation of perhaps a factor of 10^{20}, and use Eq. (6). Secondly, we may assume a linear dependence and carry out the same extrapolation with a resultant difference of a factor of 10^3 (fewer vacancies). Finally, in a fashion equivalent to the linear extrapolation, we may estimate the standard molar entropy change appropriate to the removal of a Cl atom to the vapor as half the standard molar entropy change of the reaction.

$$\text{KCl}_{(s)} = \text{K}_{(g)} + \text{Cl}_{(g)}, \quad (13)$$

plus a small contribution due to the presence of the vacancies,[2,3,16] which we somewhat arbitrarily estimate as ~4 e.u.[16] On the basis of our estimate of the total entropy change, we use Eq. (7) and the data mentioned above to calculate ΔH_{Cl}. Use of this calculated value in Eq. (9) leads, as expected, to essentially the same result

[4] J. E. Eby, K. J. Teegarden, and D. B. Dutton, Phys. Rev. 116, 1099 (1959).
[5] R. H. Bube and E. L. Lind, Phys. Rev. 105, 1711 (1957).
[6] D. Larach, R. E. Shrader, and C. F. Stocker, Phys. Rev. 108, 587 (1957).
[7] H. Tubota and H. Suzuki, J. Phys. Soc. Japan 16, 1038 (1961).
[8] D. DeNobel, thesis, University of Leiden, 1958 (unpublished).
[9] M. D. Sturge, Phys. Rev. 127, 768 (1962).
[10] J. C. Gravitt, G. E. Gross, D. K. Benson, and A. B. Scott, J. Chem. Phys. 37, 2783 (1962).
[11] M. Aven and B. Segall, Phys. Rev. 130, 81 (1963).
[12] M. R. Lorenz and B. Segall, Phys. Letters 7, 18 (1963).
[13] O. Kubaschewski and E. L. Evans, *Metallurgical Thermochemistry* (Pergamon Press, New York, 1958).
[14] D. R. Stull and G. C. Sinke, Advan. Chem. Ser. 18 (1956).
[15] C. Z. van Doorn, Phillips Res. Rept. Suppl. 4, 23 (1962).
[16] R. A. Swalin, Phys. Chem. Solids 18, 290 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

as the linear extrapolation, and is probably the most reliable method.

No thermodynamic data on vacancy formation is available for either ZnTe or GaAs. We may, however, follow Swalin[16] and write

$$\Delta H_M = \Delta H_{MN} - E_r \simeq 2E_{\text{coh}} - E_r, \qquad (14)$$

where ΔH_{MN} is the standard molar enthalpy change for the reaction

$$MN_{(s)} = M_{(g)} + N_{(g)}. \qquad (15)$$

E_{coh} is the cohesive energy per gram-atom and E_r is a "relaxation" energy associated with the vacancy. Eq. (14) is based on the assumption of constancy of bond energy and may be expected to apply to covalent crystals. An equation of the same form can actually also be derived for ionic crystals.[17] If we put the data for n-CdTe[8] into the form of Eq. (14), we find $E_r/\Delta H_{MN} = 0.24$. Similarly, for Ge,[15] $E_r/\Delta H_{MN} = E_r/2E_{\text{coh}} = 0.19$. We will assume that Eq. (14) applies to both ZnTe and GaAs. Further, we will assume that $E_r/\Delta H_{MN}$ has the same value for ZnTe as for CdTe. For GaAs, we will assume a value of $E_r/\Delta H_{MN}$ midway between that of CdTe and Ge, i.e., $E_r/\Delta H_{MN} = 0.21$, and use the data of ΔH_{GaAs}.[18]

The enthalpies and entropies of vaporization [Eq. (8)], are available for all the elements involved.[14]

For each case, we take the temperature, T, to be several hundred degrees centigrade below the maximum melting point of the host crystal in order that the maximum value of p_M be reasonably accurately described by $p_M{}^{\text{vap}}$.

The entropies of vacancy formation, ΔS_M, are generally unknown. The over-all entropy term in Eq. (9), $(\Delta S_M - \Delta S_M{}^{\text{vap}})$, is clearly the standard molar entropy change that occurs upon transfer of one mole of M atoms from $MN_{(s)}$ to $M_{(l)}$, leaving M vacancies in the crystal. This must be given by approximately minus half the standard molar entropy of formation of $MN_{(s)}$ from the solid elements ($\sim$2 e.u.) plus the standard molar entropy of fusion of $M_{(s)}$ ($\sim$3 e.u. for most elements) plus a contribution due to the "excess" entropy of the vacancies[2,3,16] (which is $\sim$4 e.u.).[15] We therefore take $(\Delta S_M - \Delta S_M{}^{\text{vap}})$ as $\sim$9 e.u. in each case.

The degeneracy factor, g_M, is taken as 2 for all cases and the concentration of available M sites, N_M, is calculated from the densities of the host crystals, which are easily available.

The density of states, N_c, is calculated from the temperature, T, and the appropriate effective mass, m^*, which is available for n-CdTe,[19,20,8] p-CdTe[8] (with some uncertainty) and n-GaAs.[21] We will take m^* for n-ZnTe as identical to that of n-CdTe and that for p-KCl as equal to the free-electron mass.

Finally, the only implicit temperature dependence in Eq. (9) that we consider is that of the bandgap, E_g. Estimates of this dependence are available for CdTe[8,22] and CdAs.[22] The value for ZnTe is interpolated from data for CdTe, CdSe, and ZnSe.[22] The value of $(\partial E_g/\partial T)$ for KCl is arbitrarily estimated as -10^{-3} eV/°C.

The inputs for Eq. (9) are given in Table I, excluding $(\Delta S_M - \Delta S_M{}^{\text{vap}}) = 9$ e.u. $= 3.9\times10^{-4}$ eV/°C and $g_M = 2$.

The results of the calculation based on Eq. (9) and the data in Table I are presented in Table II, along with data on E_g and E_{coh}. It is clear from the above discussion relating to the data in Table I that the results in the second and third columns in Table II are reliable only as to order of magnitude. Nevertheless, we may draw certain firm conclusions from these results:

1. There does exist a self-compensation "boundary" such that certain compounds, falling beyond the "boundary," cannot be doped by any equilibrium process to have appreciable electronic conductivity, e.g., the alkali halides. A rough rule of thumb for determining the position of the self-compensation "boundary" is given by the ratio of the bandgap, E_g, to the cohesive energy per gram-atom, E_{coh}; essentially complete self-compensation occurring for values of this ratio much above unity while little self-compensation occurs for values below 0.5. The expected trend from ionic to covalent compounds is observed.

2. The II-VI compounds constitute a class in which considerable self-compensation may be expected to occur ($E_g/E_{\text{coh}} \approx 1$). Nevertheless, the calculated degree of self-compensation by singly ionizable vacancies for these compounds is *not* sufficient to explain the well-known difficulty of obtaining, for example, appreciable conductivity in n-ZnTe or p-CdS. Similar results are expected for compensation by simple singly ionizable interstitial atoms [see discussion of data in Table I for p-CdTe and Eq. (14)]. We conclude, therefore, that some other limitation on the accessible carrier concentration and conductivity in II-VI compounds must be involved. There may, of course, be *impurity* limitations related to solubility or position of energy levels within the electronic bandgap. There may also be other "natural defects" involved, e.g., impurity-vacancy pairs, etc. A particular possibility that we will consider in more detail is the presence of doubly ionizable vacancies, of which both ionization levels fall within the electronic bandgap.

IV. MULTIPLY IONIZABLE VACANCIES

In the general case, we designate the concentration of M vacancies, which have been ionized to a charge of

[17] N. F. Mott and M. J. Littleton, Trans. Faraday Soc. 34, 485 (1938).
[18] V. J. Lyons and V. J. Silvestri, J. Phys. Chem. 65, 1275 (1961).
[19] D. T. F. Marple, Phys. Rev. 129, 2466 (1963).
[20] B. Segall, M. R. Lorenz, and R. E. Halstead, Phys. Rev. 129, 2471 (1963).
[21] C. Hilsum and A. C. Rose-Innes, *Semiconducting III-V Compounds* (Pergamon Press, New York, 1961).
[22] R. H. Bube, *Photoconductivity of Solids* (John Wiley & Sons, Inc., New York, 1960), p. 237.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Inputs for Eq. (9).

Compound	T (°K)	$E_g(T)$ (eV)	E_A' (eV)	ΔH_M (eV)	$\Delta H_M{}^{vap}$ (298°K) (eV)	N_M (cm^{-3})	N_c (cm^{-3})
p-KCl	970	8.03[a,c]	2.40[h]	6.20[a]	1.36[b,m]	1.38×10^{22}	1.50×10^{20} [a]
n-ZnTe	1500	1.36[d,e,a]	0.05[i]	3.43[a]	1.21[m]	1.83×10^{22}	1.03×10^{19} [a]
n-CdTe	1250	1.02[f]	0.15[f]	3.20[f]	1.07[m]	1.55×10^{22}	7.76×10^{18} [f]
p-CdTe	1250	1.02[f]	0.02[f]	3.37[a]	1.86[m]	1.55×10^{22}	3.49×10^{19} [f,k,l]
n-GaAs	1250	0.95[g,e]	0.00[a]	5.81[a,j]	2.76[m]	1.01×10^{22}	4.12×10^{18}

[a] See text in Sec. III.
[b] T is above the critical temperature of Cl_2 so that $Cl_{2(s)}$ is the condensed phase of interest rather than $Cl_{2(l)}$. We will neglect the small effect of this however.
[c] See Ref. 4. [d] See Refs. 5–7. [e] See Ref. 20. [f] See Ref. 8. [g] See Ref. 9. [h] See Ref. 10. [i] See Ref. 11. [j] See Ref. 16. [k] See Ref. 17. [l] See Ref. 18. [m] See Ref. 13.

TABLE II. Results from Eq. (9).

Compound	$Q_{calc}{}^{min}$	$\left(\frac{n}{D^+}\right)^{max}$ or $\left(\frac{p}{A^-}\right)^{max}_{calc}$	E_g (298°K) (eV)	$\Delta H_{MN}=2E_{coh}$ (eV)	E_g/E_{coh}
p-KCl	2.12×10^{8}	4.71×10^{-9}	8.70[a]	6.70[e]	2.61
n-ZnTe	3.09×10^{2}	3.22×10^{-3}	2.20[b]	4.51[e]	0.97
n-CdTe	3.23×10	3.00×10^{-2}	1.50[c]	4.20[e]	0.72
p-CdTe	8.70×10^{2}	1.15×10^{-3}	1.50[c]	4.20[e]	0.71
n-GaAs	1.70×10^{-3}	0.998	1.43[d]	7.35[f]	0.39
CdS					0.94
ZnS					1.22

[a] See Ref. 4. [b] See Refs. 5–7. [c] See Ref. 8. [d] See Ref. 9. [e] See Refs. 12 and 13. [f] See Refs. 16 and 13.

$-i$ as $V_M{}^{-i}$ and follow Sec. II in writing a charge neutrality equation,

$$n+\sum_{i=1}^{N} iV_M{}^{-i}=D^+, \tag{16}$$

where N is the highest charge state to fall within the bandgap. We also write an equilibrium constant for ionization of a vacancy in the ith charge state,

$$nV_M{}^{-i+1}/N_cV_M{}^{-i}=g_{i-1}/g_i\exp(E_{A_i}-E_c/kT), \tag{17}$$

where g_i is a degeneracy factor and E_{A_i} is an energy level describing the ith charge state of the vacancy. As before, E_c is the energy at the bottom of the conduction band. It follows from use of Eq. (17) in Eq. (16) that Eq. (5) still applies, with Q given by

$$Q=Q_1\left\{1+\sum_{i=2}^{N}(ig_i/g_1)\left(\frac{n}{N_c}\right)^{i-1}\times\exp\left[(1/kT)\sum_{j=2}^{i}(E_c-E_{A_j})\right]\right\}, \tag{18}$$

where

$$Q_1=g_1V_M/N_c\exp(E_c-E_{A_1})/kT, \tag{19}$$

as in Eq. (6).

For the II-VI compounds, it is believed that $N=2$ and we write, for these compounds,

$$Q=Q_1\{1+(2g_2/g_1)(n/N_c)\exp(E_c-E_{A_2}/kT)\}. \tag{20}$$

Certain general features of Eqs. (18) and (20) should be noted. We see that the effect of the second ionization level of the vacancy only becomes significant in Eq. (20) if the doping level is such that the Fermi level rises above the second ionization energy level [see Eq. (2)]. If we attempt to dope the host crystal so heavily that the Fermi level does rise above the second ionization level of the acceptor vacancy, we see from Eqs. (1) and (20) that the situation is soon reached in which

$$n/N_c=\exp(E_F-E_c)/kT=(g_1D^+/2g_2Q_1N_c)^{1/2}\times\exp(E_{A_2}-E_c/2kT). \tag{21}$$

It is clear that $Q_1\gg1$ for the higher bandgap II-VI compounds (see Table II). Further $(g_1D^+/2g_2Q_1N_c)^{1/2}\lesssim1$. It follows that the Fermi level in the II-VI compounds cannot be pushed closer to the bottom of the conduction band than half the energy separation between the second ionization level of the acceptor vacancy and the bottom of the conduction band. In other words, shallow impurities will be completely compensated by a combination of singly and doubly ionized vacancies in a II-VI compound if the second ionization level of the compensating vacancy falls within the bandgap.

Even if the Fermi level does not rise above the second ionization level of the acceptor vacancies, these levels will still contribute to the temperature dependence of the carrier concentration. If the donor impurities are sufficiently shallow, the position of these second ionization levels will entirely determine this temperature dependence. However, as long as $Q_1\gg1$, "freeze-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

out" of carriers onto these levels cannot push the Fermi level higher than the maximum value discussed above. Complete compensation of impurities by a combination of singly and doubly ionized vacancies will then always occur as long as the energy level of the added donor impurities is shallower (closer to the conduction band) than the second ionization level of the vacancies. If the energy level of the donor impurity is sufficiently deep, the donors will not be completely compensated, but cannot then contribute appreciable conductivity.

V. SPECIFIC IMPLICATIONS: GaAs, CdTe

The magnitude of the concentrations of vacancies in GaAs is of interest because it has been suggested that the concentration of some natural defects can be quite high ($\sim 10^{19}/\text{cm}^3$).[23] The estimated value of the energy of vacancy formation used in Sec. III, however, leads to the conclusion that the total concentration of simple vacancies, neutral or ionized, must be low. In particular, the maximum concentration of neutral Ga vacancies (under the maximum pressure of As) is calculated to be $\sim 10^{14}/\text{cm}^3$. If the Ga vacancies were *shallow* acceptors, the total concentration of Ga vacancies could hardly exceed $10^{18}/\text{cm}^3$ in a crystal of GaAs grown from an As-rich melt and doped to be degenerately n-type (at the growth or equilibration temperature). This view is supported by the well-known fact that GaAs can be prepared either n- or p-type at the $10^{17}/\text{cm}^3$ level without appreciable compensation. Thus under "normal" preparation conditions, the concentrations of electrically active natural defects must be low, i.e., $\lesssim 10^{15}/\text{cm}^3$.

As regards CdTe, it has been suggested that the second ionization level of the acceptor Cd vacancy is 0.6 eV below the conduction band.[8,12] It is, however, well known that CdTe can be doped degenerately n-type.[8,20] From the results of Sec. IV, it is clear that this contradicts the assignment of the 0.6-eV level to the second ionization level of the isolated Cd vacancy. This conclusion does not depend on the results of Sec. III but simply requires that Q_1 not be much smaller than unity. The experimental observation that substantial self-compensation is observed in n-CdTe[8] demonstrates that this requirement is met. There is, in fact, some support for the estimated energy of vacancy formation in CdTe used in Sec. III in that large concentrations of singly ionized Cd vacancies can be introduced into p-CdTe.[12]

It would thus appear that the 0.6-eV level must be assigned to some other center, possibly some complex involving a doubly ionized Cd vacancy.[12] The second ionization level of the isolated Cd vacancy in CdTe must, in fact, be much closer to the conduction band and may be the level recently observed by Lorenz *et al.*[24,25] in CdTe and some other II-VI compounds (all of which can be doped heavily n type). The exclusion of ZnTe (which apparently cannot be doped n type) from this group is significant.

VI. SOLUBILITY EFFECTS

Inasmuch as an impurity must reside on either an M site or an N site in $MN_{(s)}$, we expect that the solubility of impurities in binary compounds must, in general, be a function of the stoichiometry of the system, i.e., the concentration of vacancies. In Sec. II, we have a situation in which the concentration of added donor impurities in $MN_{(s)}$ is held fixed. Under these circumstances, it was pointed out, the maximum *relative* amount of uncompensated donor impurities will occur in a system in which the concentration of compensating acceptor vacancies is a minimum. What is generally held constant in the laboratory, however, is, to a first approximation, the thermodynamic activity of the added donor impurities. As a consequence of the dependence of the solubility of the impurities on the vacancy concentration, it is not necessarily true, then, that the maximum *total* concentration of uncompensated impurities will occur in a system in which the concentration of acceptor vacancies is a minimum.

For, example, let us consider a compound semiconductor $MN_{(s)}$, which is brought into equilibrium at the temperature T with a total constant activity of added donor impurities, i.e., the partial pressure of atomic $D_{(g)}$ is held constant at a value p_D. These impurities are presumed to reside on M sites in the host crystal for this system. The pressure of $M_{(g)}$ is again designated p_M and we assume, as in Sec. II, that only singly ionizable acceptor M vacancies need be taken into account.

The concentration of neutral donors, D, in the host crystal is determined by the equilibrium constant.[1–3]

$$D/p_D V_M = K(T), \tag{22}$$

which refers to the process in which an atom of $D_{(g)}$ "fills" an M vacancy to yield a neutral donor impurity in the lattice.

The concentration of ionized impurity atoms in the crystal, D^+, is determined by the equilibrium constant for the ionization of a neutral impurity to yield a free electron.

$$nD^+/N_c D = (1/g_D)\exp(E_D - E_c/kT), \tag{23}$$

where g_D is a degeneracy factor and E_D is the energy level associated with the donor impurity.

It follows, from Eqs. (5) and (23), that

$$n^2 = [1/(1+Q)](N_c D/g_D)\exp(E_D - E_c)/kT). \tag{24}$$

At constant p_D and T, D is proportional to V_M according to Eq. (22). On the other hand, according to

[23] J. Blanc, R. H. Bube, and L. R. Weisberg, Phys. Rev. Letters 9, 252 (1962).

[24] M. R. Lorenz and H. H. Woodbury, Phys. Rev. Letters 10, 215 (1963).

[25] M. R. Lorenz, M. Aven, and H. H. Woodbury, Phys. Rev. 132, 143 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Eq. (6), Q is also proportional to V_M at constant T. We see, therefore, that the total concentration of free carriers, n, is a maximum when V_M is a *maximum*, despite the fact that the relative amount of uncompensated donors, n/D^+, is a minimum under the same conditions. When $Q \gg 1$, as in the II-VI compounds (see Sec. III), the *total* concentration of free carriers is essentially independent of V_M. Thus, as p_M is varied in these systems, only the concentration of compensated donors varies.

It should be noted that the situation is entirely different if the donor impurities reside on N sites rather than M sites. Under such circumstances, the solubility of the impurities is maximized and the degree of self-compensation minimized when V_M is minimized, i.e., when p_M is maximized.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Theory of the Optical and Magnetic Properties of the Self-Trapped Hole in Lithium Fluoride*

T. P. Das and A. N. Jette
University of California, Riverside, California

and

R. S. Knox
University of Rochester, Rochester, New York

(Received 23 September 1963)

Using a semiphenomenological method, the energy and wave functions of a self-trapped hole (V_k center) in LiF are obtained as a function of the separation between the two F^- ions at which the hole is assumed trapped. The lattice distortion energy due to the changes in Madelung, repulsive, and polarization energies is calculated as a function of the totally symmetric displacement of the two participating F^- ions and six positive ions adjacent to the F^- ions. This lattice energy is combined with the calculated energy for the F_2^- molecule to obtain the total energy as a function of the distance between the participating F^- ions for both the symmetric (Σ_g) and antisymmetric (Σ_u) states of the hole on the V_k center. Only the energy curve for the ground (Σ_u) state exhibits a minimum in the expected region of F^--ion separation. From the resulting configurational coordinate curves, the optical absorption energy and width are computed and found to be in order-of-magnitude agreement with experiment. Computed values of the experimentally known isotropic and anisotropic hyperfine constants are used to assess the validity of our molecular wave functions, which were obtained in a one-electron approximation.

I. INTRODUCTION

In recent years, a fairly detailed understanding of the F center in alkali halides has been achieved by concerted theoretical and experimental studies of its electronic structure.[1] The F center consists of an electron bound to a negative-ion vacancy, and is the simplest of several "electron" color centers such as the M, R, and F' centers.[2] A series of "hole" color centers also exists (V_1,H,V_k); these are characterized by optical absorption bands lying at somewhat higher energies than those of electron centers, are generally stable only at low temperatures, and must be formed by high-energy irradiation. The best understood of these, a center which consists of a self-trapped hole, has been studied elaborately by Castner and Känzig.[3] Electron resonance data have conclusively shown that this center is not associated with a vacancy but instead resembles a negative halogen molecule-ion with a hole shared between two adjacent negative ions. Two previous theoretical attempts[4,5] have been made to explain the stability of this self-trapped holes. Yamashita[4] tried to calculate the energy of the hole in KCl as a function of the displacement of the Cl^- ions which trap it. He could not obtain a minimum in the energy as a function of this displacement but suggested that the repulsion between the core electrons of the Cl^- ions, which he had neglected, might provide a minimum in the energy curve at half the usual distance between the ions in the crystal. A subsequent attempt by Nettel[5] gave the result that the energy of a hole trapped as a Cl_2^- ion in the crystal had higher energy than a chlorine atom

* Supported in part by the National Science Foundation and in part by a Grant (62–145) from the U. S. Air Force Office of Scientific Research.

[1] See, for example, the review article by B. S. Gourary and F. J. Adrian in *Solid State Physics* edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1960), Vol. 10.

[2] See, for example, J. H. Schulman and W. D. Compton, *Color Centers in Solids* (Pergamon Press, Inc., New York, 1963).

[3] T. G. Castner and W. Känzig, Phys. Chem. Solids **3**, 178 (1957); Nuovo Cimento Suppl. **7**, 612 (1958).

[4] J. Yamashita, work at the University of Illinois, 1958 (unpublished).

[5] S. J. Nettel, Phys. Rev. **121**, 425 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

at a cubic site. It appears that Nettel had not employed adequate variational functions for the hole to take account of the relatively strong distortion of the wave function of one of the trapping ions by the other.

In the present paper we attempt to obtain the energy of the LiF crystal containing a hole relative to that of the perfect crystal. The hole is assumed to be shared between two F^- ions so that we have in effect a F_2^- molecule. The energies of the three σ electrons are calculated as a function of the distance between the ions in the two cases corresponding to the unpaired electron being in a symmetric or antisymmetric state. For each distance between the ions, the repulsive, Madelung, and polarization energies are also calculated allowing for relaxation and electronic polarization of six positive ions in the plane of the F_2^- molecule, which are the nearest neighbors to the two F^- ions forming the molecule. The minimum in the total energy curve plotted as a function of the distance between the fluorine nuclei of the F_2^- molecule then gives the equilibrium configuration of the crystal.

In Sec. II, the various contributions to the energy are discussed qualitatively. In Sec. III, the calculation of the lattice energy change is described and in Sec. IV, the computations for obtaining the electronic energies and wave function of the F_2^- molecule are described. In Sec. V, the configuration coordinate curves for the crystal are presented and the optical absorption energy and width and hyperfine constants are calculated and compared with experiment. The nature of the agreement between theoretical and experimental results is then analyzed. In the last section, limitations of the present work are discussed and suggestions for improvement are listed.

II. ANALYSIS OF THE VARIOUS TERMS IN THE ENERGY OF THE CRYSTAL

Our calculation is based on the variation principle. The energy of the crystal containing the hole is calculated as a function of several variation parameters. Some of these parameters may be regarded as equivalent to the configuration coordinates in calculations on luminescent centers, while others are variation parameters that occur in the electronic wave functions. With respect to the former, our calculation resembles that of Williams,[6] with at least one major exception. He was able to use experimental values of the excitation energy of free Tl^+ ion in constructing his excited state configurational coordinate curves, whereas we must rely on computations of this quantity.

Before we describe our variation parameters in detail, we will list and explain the various terms in the total crystal energy E when a hole is present

$$E=E_0+E_A+(\alpha_M e^2/a)+E_E+\Delta E_M+\Delta E_R+\Delta E_P. \quad (1)$$

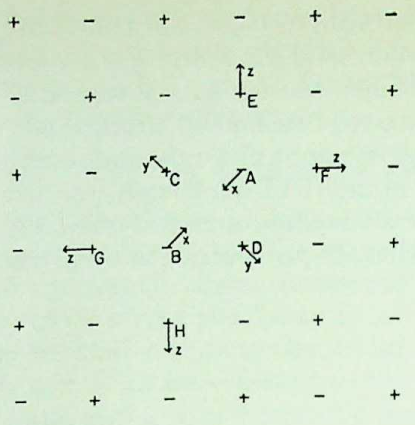


FIG. 1. Diagram showing the ions $A, B, \cdots, H$ which we allow to displace and their directions of displacement.

These terms may be understood in terms of a sequence of events occurring during the removal of an electron from the crystal. E_0 is the energy of the perfect crystal; the next two terms represent the energy required to remove an electron from an F^- ion and place it at infinity; E_A is the electron affinity (work done in taking the electron from F^- itself) and $\alpha_M e^2/a$ is the Madelung energy (work done taking the electron out of the electrostatic potential of the crystal). Here a is the nearest-neighbor distance and α_M the corresponding Madelung constant. Up to this point we have regarded the lattice and remaining electrons as stationary; now let us assume that the hole becomes shared with a nearest neighboring F^- ion so that a molecule F_2^- is formed (see ions A and B in Fig. 1). This will "require" an energy E_E, a quantity which is negative and which will certainly depend on R, the nuclear separation of A and B. The redistribution of charge on the two F^- ions, as well as their relaxation toward each other, will cause a further change in electrostatic energy of the crystal which is included in ΔE_M. In this term is also included the change of electrostatic energy due to the relaxation of neighboring positive ions which will now occur as explained in detail below. Finally, ΔE_R and ΔE_P are the analogous changes in repulsive and polarization energies, respectively, occasioned by the redistribution of charge and motion of the ions.

Of the six terms in (1) which represent the change in the energy of the crystal due to the creation and capture of a hole, the last three and $\alpha_M e^2/a$ may be considered as constituting a change in lattice energy $\Delta E_L=\Delta E_M+\Delta E_P+\Delta E_R+(\alpha_M e^2/a)$. These three contributions are calculated as a function of the fractional displacements (in terms of nearest-neighbor distance in the lattice) x, y, and z of the ions of types A, C, and E shown in Fig. 1. The type-A ions are the two F^- ions (A,B) which share the hole, type C are the two Li^+ ions (C,D) which are the nearest neighbors of A and B. The type E ions are the four Li^+ ions (E,F,G,H) which are the nearest neighbors of the A and B ions. They lie on the same plane as the A- and C-type ions. The relaxation of the four Li^+ ions, which are nearest neighbors of the A and B ions but which do not lie on the plane of the A- and C-type ions, is neglected. Because of the geometry of

[6] F. E. Williams, J. Chem. Phys. 19, 457 (1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the location of these ions, they are an order of magnitude less effective in influencing the lattice energy than the ions C, D, E, F, G, and H. In calculating the repulsive energy change ΔE_R, only nearest-neighbor interactions are considered. The polarization energy of the rest of the lattice as well as of the ions $A, \cdots, F$ is calculated following Mott and Littleton's procedure.[7] The net lattice energy is thus obtained as a function $\Delta E_L(x,y,z)$. Since x represents the major relaxation, $\Delta E_L(x,y,z)$ is minimized with respect to y and z to obtain $\Delta E_L(x,y_0,z_0)$ which is a function of x alone. The details and results of the lattice energy calculation are presented in Sec. III.

The terms E_A and E_E represent the electronic energy changes associated with the creation and capture of the hole. The electron-affinity E_A of the F^- ion is taken as 0.363 eV, from Hartree-Fock calculations[8] on the F^- ion. The F_2^- binding energy E_E is obtained by calculating the energy of three σ electrons in the molecule (two electrons in σ_g state and one in σ_u when the hole is in the antisymmetric state Σ_u, and two electrons in σ_u state and one in σ_g state for the hole in the symmetric state Σ_g) and taking the difference between this energy and the energy of three $2p$ electrons on the F^- ion. The details of the calculation of E_E as a function of the distance between A and B, that is, $E_E(x)$, as well as the potential and the variational form of the wave function employed will be discussed in Sec. IV. $E_E(x)$ is then combined with $\Delta E_L(x)$ to obtain the variation of the total energy of the crystal with x; the position of its minimum gives the equilibrium configuration of the crystal. It will be seen later that a minimum is found only when the hole is in the Σ_u and not in the state Σ_g.

III. CALCULATION OF LATTICE ENERGY CHANGES

We shall first consider the change in the Madelung energy of the lattice due to the presence of the hole. For this calculation, we shall assume that the hole is equally divided between the two F^- ions. This is a fairly good approximation if the bonding between A and B is not too strong so that there is not much migration of charge to the region between the two ions. The general expression for the difference in Madelung energies of the crystal between the two situations in which the ions $A, B, C, \cdots, H$ are displaced, $E_M(x,y,z)$, and undisplaced, $E_M(0,0,0)$, can be shown to be given by

$$E_M(x,y,z)-E_M(0,0,0)$$
$$=\sum_{L=A}^{H}\sum_{M\neq A,\cdots,H}^{\infty}\zeta_L\zeta_M\left(\frac{1}{r_{LM_0}}-\frac{1}{r_{L_0M_0}}\right)$$
$$+\sum_{L=A}^{M-1}\sum_{M=A}^{H}\zeta_L\zeta_M\left(\frac{1}{r_{LM}}-\frac{1}{r_{L_0M_0}}\right). \quad (2)$$

Here ζ_L and ζ_M are the charges on the lattice points L and M, and L_0 and L represent, respectively, the undisplaced and displaced positions of the point L. Thus, for example, r_{LM_0} represents the distance between the displaced position of the point L and undisplaced position of M. If we are interested in an expansion of (2) in powers of x, y, and z correct to second order, we may take

$$\sum_{M\neq L}^{\infty}\frac{\zeta_L\zeta_M}{r_{LM_0}}=\sum_{M\neq L}^{\infty}\frac{\zeta_L\zeta_M}{r_{L_0M_0}}, \quad (3)$$

provided that all ζ_M are equal to $\pm e$ and are situated on a regular cubic lattice. Equation (3) is a result of the fact that in regular cubic crystals, the next multipole interaction between ions after the monopole corresponds to $l=4$. Hence only quartic and higher terms in x, y, or z are to be expected from the summation in (3). It follows that

$$\sum_{M\neq A\cdots H}^{\infty}\frac{\zeta_L\zeta_M}{r_{LM_0}}=\sum_{M\neq L}^{\infty}\frac{\zeta_L\zeta_M}{r_{L_0M_0}}-\sum_{\substack{M=A\\M\neq L}}^{H}\frac{\zeta_L\zeta_M}{r_{LM_0}}, \quad (4)$$

which may be used to simplify Eq. (2) to the following form:

$$E_M(x,y,z)-E_M(0,0,0)$$
$$=\frac{1}{2}\sum_{\substack{M=A\\M\neq L}}^{H}\sum_{L=A}^{H}\zeta_L\zeta_M\left(\frac{1}{r_{LM}}-\frac{2}{r_{LM_0}}+\frac{1}{r_{L_0M_0}}\right). \quad (5)$$

If the charges on the A and B ions are $-qe$ and the charges on the rest of the lattice points are the same as in the perfect lattice, then the above calculation must be modified in an obvious way, and one obtains

$$E_M(x,y,z)-E_M(0,0,0)$$
$$=\sum_{L=A}^{M-1}\sum_{M=A}^{H}\zeta_L\zeta_M\left(\frac{1}{r_{LM}}+\frac{1}{r_{L_0M_0}}-\frac{1}{r_{LM_0}}-\frac{1}{r_{L_0M}}\right)$$
$$+2(q^2-q)\left(\frac{1}{r_{AB_0}}-\frac{1}{r_{A_0B_0}}\right)$$
$$-2(q-1)\sum_{L=C}^{H}\left(\frac{1}{r_{LA_0}}-\frac{1}{r_{L_0A_0}}\right). \quad (6)$$

In the present case we are assuming that $q=\frac{1}{2}$. Also, ζ_L is $+e$ for a positive ion and $-e$ for a negative ion. With the help of Fig. 1, one can calculate the various distances involved in Eq. (6) correct to terms quadratic in x, y, and z. Thus, for example,

$$\frac{1}{r_{AE}}+\frac{1}{r_{A_0E_0}}-\frac{1}{r_{AE_0}}-\frac{1}{r_{A_0E}}=-\frac{\sqrt{2}}{a}xz. \quad (7)$$

Equation (6) includes the repulsive energy between charges $-e/2$ on the ions A and B. This is a spurious

[7] N. F. Mott and M. J. Littleton, Trans. Faraday Soc. 34 485, (1938).
[8] C. Froese, Proc. Cambridge Phil. Soc. 53, 206 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

term introduced by our assumption that the hole is equally distributed between the two F^- ions, and must be excluded because all such intramolecular interactions are contained in E_E. We therefore obtain the following expression for ΔE_M:

$$\Delta E_M(x,y,z)=E_M(x,y,z)-E_M(0,0,0)-[e^2/4\sqrt{2}a(1-\sqrt{2}x)]. \quad (8)$$

Using equations like (7), Eqs. (6) and (8) may be expanded in powers of x, y, and z, and after a lengthy calculation one obtains

$$\Delta E_M=(e^2/a)(-0.177-0.500x-0.354x^2-1.414y+1.207y^2+2.358z+3.426z^2+3.000xy+2.500zx-1.707yz). \quad (9)$$

It should be emphasized that this equation has been derived for a particular (fcc) lattice.

Next, we consider ΔE_R. In keeping with our earlier assumptions for the Madelung energy change we shall not consider the repulsion between the ions A and B. This latter repulsive energy, when it becomes appreciable, will be included in the electronic energy terms. Also, we shall consider repulsion between nearest neighbors alone. There are several choices that one can make for the repulsive interaction energy between the positive and negative ions. One possible choice is the Born-Mayer potential,[9] which has the form

$$V(r)=Be^{-r/\rho}, \quad (10)$$

with B and ρ obtained from the observed compressibility and lattice constant of the crystal. In the determination of B and ρ in the Born-Mayer formula from the perfect crystal, Van der Waals forces were neglected, i.e., their effect is assumed to be adequately included in the exponential parameters. Therefore in using this formula for the imperfect crystal, we should ignore Van der Waals forces for the sake of consistency. Huggins and Mayer[10] have also obtained B and ρ from data on the perfect crystal, but including Van der Waals forces separately. We shall also make use of these constants to test the sensitiveness of the calculated lattice energy to the choice of force constants. Another form for the repulsive energy has been proposed by Pauling,[11] namely

$$V(r)=0.0485e^2r_0{}^{n-1}/r^n, \quad (11)$$

where r_0 is the sum of the ionic radii of the positive and negative ions and n depends on the nature of the positive and negative ion pair. For LiF one can take $n=6$ so that for the repulsive potential between Li^+ and F^- ion we can rewrite (11) as

$$V(r)=Ae^2/r^6, \quad (12)$$

where $A=0.0485r_0{}^5$. Using the form in Eq. (10) for the repulsive energy, the expression for $\Delta E_R(x,y,z)$ is given by Eq. (13):

$$\Delta E_R(x,y,z)=4Be^{-a/\rho}\left[\frac{a}{\rho}\left\{\left(-x^2-y^2-2z^2-\frac{xy}{2}\right)+\frac{a}{2\rho}(x^2+y^2+2z^2-xy-\sqrt{2}xz)\right\}\right]. \quad (13)$$

In obtaining Eq. (13), the energy differences between a pair of ions L and M is calculated using the relation

$$\Delta E_R(L,M)=B[e^{-r_{LM}/\rho}-e^{-r_{L_0M_0}/\rho}], \quad (14)$$

where r_{LM} and $r_{L_0M_0}$ have the same meaning as in Eq. (6). The distances r_{LM} between the displaced ions and their nearest neighbors are calculated up to quadratic terms in x, y, and z as before.

Using the Pauling expression (12) for the force law, Eq. (15) is obtained for the change in the repulsive energy as a function of x, y, z:

$$\Delta E_R(x,y,z)=(Ae^2/a^6)[60x^2+60y^2+120z^2-96xy-84\sqrt{2}xz]. \quad (15)$$

In Table I are tabulated the constants B and ρ both in the Born-Mayer and Huggins-Mayer approximations as well as the value of A in the Pauling approximation. In obtaining A, a value of $r_0=1.95$ Å was used in Eq. (12), this value of r_0 being the sum of the ionic radii of Li^+ and F^- ions. The lattice constant a for LiF is 2.01 Å, only slightly different from r_0. On introducing the values of B, ρ, and A in Eqs. (14) and (15), we get the following equations for $\Delta E_R(x,y,z)$ in the three cases.

Huggins-Mayer:

$$\Delta E_R(x,y,z)=\frac{e^2}{a}[1.664x^2+1.664y^2+3.328z^2-3.031xy-3.642xz], \quad (16a)$$

Born-Mayer:

$$\Delta E_R(x,y,z)=\frac{e^2}{a}[2.482x^2+2.482y^2+4.964z^2-3.675xy-4.635xz], \quad (16b)$$

Pauling:

$$\Delta E_R(x,y,z)=\frac{e^2}{a}[2.910x^2+2.910y^2+5.820z^2-4.656xy-5.762xz]. \quad (16c)$$

The difference between Eqs. (16a) and (16b) which use the same force law but different force constants B and ρ could partly be ascribed to the neglect of the effect of the Van der Waals forces in using the Huggins-Mayer formula. In order to get an expression for

[9] M. Born and J. E. Mayer, Z. Physik. 75, 1 (1932); we will use the simple theory given by M. Born and K. Huang, *Dynamics of Crystal Lattices* (Oxford University Press, New York, 1955), p. 26.

[10] M. L. Huggins and J. E. Mayer, J. Chem. Phys. 1, 643 (1933).

[11] L. Pauling, Proc. Natl. Acad. Sci. India Sec. A, Pt. I 25, (1956). See also Milne and Cubiccoti, J. Chem. Phys. 29, 846 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Table of constants in repulsive energy formulas.

	B (ergs)	ρ (Å)	A
Born-Mayer	1.05×10^{-9}	0.244	
Huggins-Mayer	1.3541×10^{-10}	0.345	
Pauling			1.368×10^{-40}

$\Delta E_R(x,y,z)$ which would be comparable to (16b) one would have to include in $\Delta E_R(x,y,z)$ the contribution from the change in Van der Waals energy due to the displacement of the lattice points. We shall not do this but instead use the differences between the lattice energy changes with Huggins-Mayer and Born-Mayer formulas as a measure of the importance of Van der Waals forces. Equations (16a) and (16b) indicate that while the Van der Waals contribution to the lattice energy is significant, it is definitely less important than the repulsive energy change. The difference between the energy expressions in (16b) and (16c) is due to the different formulas [Eqs. (10) and (12)] employed. It is encouraging that the difference between Eqs. (16b) and (16c) is less than that between (16a) and (16b) since both Born-Mayer and Pauling formulas include the effect of the Van der Waals forces indirectly. One last comment about the repulsive energy calculation. We have used Eqs. (10) and (12) together with the constants in Table I which refer to the energy of repulsion between Li^+ and F^- ions. Actually ions A and B are somewhat different from the regular F^- ion because they each contain a total charge $-e/2$, which is less than that of a regular F^- ion. Our belief that the use of Eqs. (10) and (12) does not lead to significant error is based on two considerations. First, that the charge distribution of five and one-half $2p$ electrons is probably not too different from what it would be for six. Secondly, at least in the Σ_u state, the hole probably affects only the charge distribution in the region between the two ions A and B. Hence, to the nearest-neighbor positive ions, $C, D, \cdots, H$, the A and B ions do not "look" very different from perfect F^- ions.

We shall finally consider ΔE_P, the energy associated with the polarization of the lattice due both to the two holes on ions A and B and the displacements of the neighboring points $C, D, \cdots, H$. The polarization energy due to a point charge at a lattice site has been considered in detail by Mott and Littleton.[12] In the zeroth-order approximation, where one considers the crystal as a dielectric continuum in calculating local fields at the ions, the dipole moments on the positive and negative ions are given by

$$N_\pm = \frac{M_\pm a^3 e}{r^2}, \tag{17}$$

with the dipoles directed outwards in the direction of $\mathbf{r}$, the radius vector joining the impurity charge ζ_i and the ion under consideration.

$$M_\pm = \frac{\zeta_i(1-1/K_S)(e^2/A+\alpha_\pm)}{4\pi[(e^2/A)+\frac{1}{2}(\alpha_+ +\alpha_-)]}, \tag{18}$$

where α_+ and α_- are the electronic polarizabilities of the positive and negative ions on the lattice, K_S is the static dielectric constant and $A = 6a/\theta$ is the spring constant with the compressibility $1/\theta$ given by

$$\frac{1}{\theta} = \frac{1}{3a}\left[\frac{1}{\rho^2} - \frac{2}{a\rho}\right] B e^{-a/\rho}. \tag{19}$$

We however follow the first-order approximation of Mott and Littleton in that we permit the nearest-neighbor ions ($B, C, \cdots, H$ in the case of the hole on A) to move discretely and determine their displacements by a detailed consideration of the various terms in the lattice energy. So we have to allow for this difference in nature between the ions $A, \cdots, H$ and the ions in the rest of the crystal. We shall do this by taking into account only the electronic polarization of ions A to H, since their ionic polarizations have already been considered by calculating their electrostatic and repulsive interactions with other ions in the crystal in their displaced positions.

Our polarization energy calculation is also somewhat more involved than the calculation of the polarization energy due to single point charges such as single vacancies and interstitials in a lattice. The ions A and B carry only charges $-e/2$ and are in displaced positions as compared to their original positions A_0 and B_0 in the lattice. Therefore, in the calculation of polarization effects we have in effect to consider four charges, namely $-\frac{1}{2}e$ each at A and B, and $+e$ each at the points A_0 and B_0. The latter two represent the effective charges the vacancies would have if the charges $-\frac{1}{2}e$ were *entirely* removed. Each of these four charges polarizes the ions in the rest of the lattice and the dipole moment on the polarized ion not only interacts with the polarizing charge (direct polarization energy) but also with the other three charges (indirect polarization energy). We shall consider the four ions in six pairs, namely (A_0B_0), (AB), $(AB_0)(A_0B)$, (AA_0), and (BB_0). To avoid counting the direct polarization energy for each of the four charges twice, we shall include the direct polarization energies only while considering the first two pairs (A_0B_0) and (AB). Strictly speaking, one should consider similar pairs for the ions C, D, E, F, G, H because they are also displaced from their original equilibrium positions. A consideration of such effects is equivalent to considering the polarization of the lattice due to the dipoles produced by the displacements of $C, \cdots, H$ from their original positions. Since the displacements of the ions $C, \cdots, H$ are expected to be less important than those of A and B, we shall neglect the

[12] See Ref. 7; B. G. Dick (private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

effects of the dipoles at $C, \cdots, H$ on the polarization energy. Thus we have

$$\Delta E_P = \Delta E_p(A_0B_0)+\Delta E_p(AB)+\Delta E_p(AB_0)+\Delta E_p(A_0B)+\Delta E_p(AA_0)+\Delta E_p(BB_0). \tag{20}$$

$$\Delta E_p(A_0B_0) = -\frac{e^2}{a}\Bigg[M_1\sum_i \frac{1}{r_{A_0i}^4}+M_2\sum_j \frac{1}{r_{A_0j}^4}+M_1\sum_i \frac{\cos\theta_{A_0iB_0}}{r_{A_0i}^2 r_{B_0i}^2}+M_2\sum_j \frac{\cos\theta_{A_0jB_0}}{r_{A_0j}^2 r_{B_0j}^2}-\frac{M_2}{r_{A_0B_0}^4}-2\left(\frac{M_1}{r_{A_0C_0}^4}-\frac{M_1'}{r_{A_0C}^4}\right)$$
$$-2\left(\frac{M_1\cos\theta_{A_0C_0B_0}}{r_{A_0C_0}^2 r_{B_0C_0}^2}-\frac{M_1'\cos\theta_{A_0CB_0}}{r_{A_0C}^2 r_{B_0C}^2}\right)-2\left(\frac{M_1}{r_{A_0E_0}^4}-\frac{M_1'}{r_{A_0E}^4}\right)$$
$$-2\left(\frac{M_1}{r_{A_0G_0}^4}-\frac{M_1'}{r_{A_0G}^4}\right)-4\left(\frac{M_1\cos\theta_{A_0E_0B_0}}{r_{A_0E_0}^2 r_{B_0E_0}^2}-\frac{M_1'\cos\theta_{A_0EB_0}}{r_{A_0E}^2 r_{B_0E}^2}\right)\Bigg], \tag{21}$$

where

$$M_1=\frac{[1-(1/K_S)][(e^2/A)+\alpha_+]}{4\pi[(e^2/A)+\frac{1}{2}(\alpha_++\alpha_-)]}, \quad M_1'=\frac{\alpha_+}{a^3},$$
$$M_2=\frac{[1-(1/K_S)][(e^2/A)+\alpha_-]}{4\pi[(e^2/A)+\frac{1}{2}(\alpha_++\alpha_-)]}, \quad M_2'=\frac{\alpha_-}{a^3}. \tag{22}$$

The summation in i in the first and third terms of (21) runs over the positive ions in the lattice while the summation over j in the second and fourth terms runs over the negative ions except that in the fourth term the points $j=A_0$ and $j=B_0$ are excluded. The fifth term arises because (when considering the effect of the charge $+e$ at A_0) we do not have an ion at the B_0 site and vice versa for the charge $+e$ at B_0. The sixth, seventh, eighth, ninth, and tenth terms in (21) take account of the correction to be introduced because we have to consider only the electronic polarization of the displaced ions C, D, E, F, G, and H. The angles like $\theta_{A_0jB_0}$ stand for the angle subtended by the points A_0 and B_0 at the lattice point j.

$$\Delta E_p(AB) = -\frac{e^2}{4a}\Bigg[M_1\sum_i \frac{1}{r_{A_i}^4}+M_2\sum_j \frac{1}{r_{A_j}^4}+M_1\sum_i \frac{\cos\theta_{A_iB}}{r_{A_i}^2 r_{B_i}^2}+M_2\sum_j \frac{\cos\theta_{A_jB}}{r_{A_j}^2 r_{B_j}^2}-(M_2-M_2'')\frac{1}{r_{AB}^4}$$
$$-2\left(\frac{M_1}{r_{AC_0}^4}-\frac{M_1'}{r_{AC}^4}\right)-2\left(\frac{M_1\cos\theta_{AC_0B}}{r_{AC_0}^2 r_{BC_0}^2}-\frac{M_1'\cos\theta_{ACB}}{r_{AC}^2 r_{BC}^2}\right)-2\left(\frac{M_1}{r_{AE_0}^4}-\frac{M_1'}{r_{AE}^4}\right)-2\left(\frac{M_1}{r_{AG_0}^4}-\frac{M_1'}{r_{AG}^4}\right)$$
$$-4\left(\frac{M_1\cos\theta_{AE_0B}}{r_{AE_0}^2 r_{BE_0}^2}-\frac{M_1'\cos\theta_{AEB}}{r_{AE}^2 r_{BE}^2}\right)\Bigg], \tag{23}$$

where M_1, M_2, and M_1' are defined in Eq. (22) and

$$M_2''=\frac{\alpha_-'}{a^3}. \tag{24}$$

The quantity α_-' represents the polarizability of the π electrons on the F^- ion. The fifth term is introduced to take account of the fact that only the polarization of the π electrons on A and B ions has to be considered; the polarization of the σ electrons will be taken care of indirectly in the electronic energy. The summations in the first and third terms again run over the positive ions alone and those in the second and fourth terms over the negative ions alone, the points $j\neq A, B$ are again excluded from the fourth summation. The rest of the terms are introduced to avoid introducing the ionic polarization of the ions $C, D, \cdots, H$ in ΔE_P.

$$\Delta E_p(AB_0)+\Delta E_p(A_0B) = \frac{e^2}{a}\Bigg[M_1\sum_i \frac{\cos\theta_{A_0jB}}{r_{A_0i}^2 r_{Bi}^2}+M_2\sum_j \frac{\cos\theta_{A_0jB}}{r_{A_0j}^2 r_{Bj}^2}-2\left(\frac{M_1\cos\theta_{A_0C_0B}}{r_{A_0C_0}^2 r_{BC_0}^2}-\frac{M_1'\cos\theta_{A_0CB}}{r_{A_0C}^2 r_{BC}^2}\right)$$
$$-2\left(\frac{M_1\cos\theta_{A_0E_0B}}{r_{A_0E_0}^2 r_{BE_0}^2}-\frac{M_1'\cos\theta_{A_0EB}}{r_{A_0E}^2 r_{BE}^2}\right)-2\left(\frac{M_1\cos\theta_{A_0G_0B}}{r_{A_0G_0}^2 r_{BG_0}^2}-\frac{M_1'\cos\theta_{A_0GB}}{r_{A_0G}^2 r_{BG}^2}\right)\Bigg]. \tag{25}$$

In Eq. (25) also, the summation indices (in the first and second terms) refer to positive and negative ions, re-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

spectively, the ion B being left out of the summation in j. Finally, we have

$$\Delta E_p(AA_0)+\Delta E_p(BB_0)=-\frac{e^2}{a}\Bigg[M_1\sum_i\frac{\cos\theta_{A_0iA}}{r_{A_0i}^2r_{Ai}^2}+M_2\sum_j\frac{\cos\theta_{A_0jA}}{r_{A_0j}^2r_{Aj}^2}-2\left(\frac{M_1\cos\theta_{A_0C_0A}}{r_{A_0C_0}^2r_{AC_0}^2}-\frac{M_1'\cos\theta_{A_0CA}}{r_{A_0C}^2r_{AC}^2}\right)$$
$$-2\left(\frac{M_1\cos\theta_{A_0E_0A}}{r_{A_0E_0}^2r_{AE_0}^2}-\frac{M_1'\cos\theta_{A_0EA}}{r_{A_0E}^2r_{AE}^2}\right)-2\left(\frac{M_1\cos\theta_{A_0G_0A}}{r_{A_0G_0}^2r_{AG_0}^2}-\frac{M_1'\cos\theta_{A_0GA}}{r_{A_0G}^2r_{AG}^2}\right)\Bigg]. \quad (26)$$

We can combine Eqs. (21), (23), (25), and (26) to obtain the net polarization energy of the lattice given by Eq. (20). Since we require the energy only up to quadratic terms we have to express the polarization energy in terms of linear and quadratic terms in x, y, and z only. This can be done with the help of Fig. 1 and expressing the distances and angles involving the displaced ions A, B, $\cdots$, H in terms of linear and quadratic terms in x, y, and z. The detailed evaluation of the coefficients of these terms requires several lattice summations, some of them about the single center A and some involving the two centers A and B. A few of these summations were available from the work of earlier authors,[7,13] namely

$$\sum_i\frac{1}{r_{A_0i}^4}=10.1977,$$
$$\sum_i\frac{1}{r_{A_0j}^4}=6.3346,$$
$$\sum_i\frac{1}{r_{A_0i}r_{B_0i}^3}-\sum_i\frac{1}{r_{A_0i}^3r_{B_0i}^3}=5.036, \quad (27)$$
$$\sum_i\frac{1}{r_{A_0j}r_{B_0j}^3}-\sum_i\frac{1}{r_{A_0j}^3r_{B_0j}^3}=4.152.$$

These summations are the least convergent of the various sums needed. The rest of the sums have the general form

$$\sum\frac{x_{A_0i}^{\alpha}x_{B_0i}^{\beta}y_{A_0i}^{\gamma}y_{B0i}^{\delta}}{r_{A_0i}^{\sigma}r_{B_0^{\sigma}i}^{\tau}}, \quad (28)$$

with similar expressions involving j. x_{A_0i}, y_{A_0i}, and z_{A_0i} represent coordinates of the ith ion in terms of the crystal axes with respect to the undisplaced position of A. These were performed on the Illiac at the University of Illinois computing center and checked on the IBM-1620 at the computing center of the Riverside campus of the University of California. Their values are presented in Table II. The sums were carried out in all cases up to the eleventh neighbors of the relevant type (positive or negative ion), and the convergence was tested by finding the additional contribution in going to the thirteenth neighbors. The convergence in all cases was better than 0.1%.

Next we have to consider the values of the quantities M_1, M_2, M_1', M_2'' to be used in the calculation. The expressions for these are given in terms of K_S, α_+, α_-, α_-', and the spring constant A in Eqs. (22) and (24). In Table III the values of these quantities are tabulated together with the source from which they were obtained.

TABLE II. Lattice sums involved in the calculation of ΔE_p. The sums and notation are defined in the text [see Eq. (28)].

α	β	γ	δ	σ	τ	Sum over i	Sum over j
0	0	0	0	1	5	4.62961	1.32584
0	0	0	0	3	5	2.79582	0.41540
0	1	0	0	3	3	1.73987	0.54498
0	1	0	0	1	5	0.73710	0.13719
0	1	0	0	3	5	1.11161	0.14987
0	2	0	0	3	5	1.42513	0.34914
0	1	0	1	3	5	0.18422	0.00136
1	1	0	0	3	5	0.31351	0.19927
1	0	0	1	3	5	−0.92679	−0.14850
1	1	0	0	5	5	0.08821	0.04716
0	1	0	1	1	7	0.02760	−0.05133
0	2	0	0	3	7	1.17244	0.12574
0	1	0	1	3	7	0.05051	−0.01262
0	2	0	0	1	7	1.64155	0.41656
0	2	0	0	0	8	2.19260	0.59693
1	0	0	1	5	5	−0.97343	−0.07417

TABLE III. Values of the quantities involved in expressions for M_1, M_2, M_1', and M_2''.

Quantity	Value	Source
K_S	9.27	M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1954), Table 17.
α_+	0.03×10^{-24} cm³	J. R. Tessman, A. H. Kahn and W. Shockley, Phys. Rev. 92, 890 (1953).
α_-	0.652×10^{-24} cm³	C. Kittel, *Solid State Physics*, (John Wiley and Sons Inc., New York, 1959), p. 165.
α_-'	0.391×10^{-24} cm³	Represents polarizability of π electrons which is $\frac{3}{5}$ of α_- [see R. M. Sternheimer, Phys. Rev. 96, 951 (1954)].
A	1.17×10^{-12}	M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1954), Table 9.

[13] J. R. Reitz and Gammel, J. Chem. Phys. 19, 894 (1951); F. Bassani and N. Inchauspé, Phys. Rev. 105, 819 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Using these values, we obtain

$$\begin{aligned} M_1&=0.0624,\\ M_1'&=0.0036,\\ M_2&=0.0796,\\ M_2'&=0.0482. \end{aligned} \tag{29}$$

The result of our expansion and substitution is finally,

$$\begin{aligned}\Delta E_P(x,y,z)=-(e^2/4a)[&1.416-0.505x+0.702x^2\\ &-0.010y-0.007y^2+0.089xy+0.037z\\ &+0.087z^2+0.106xz]. \end{aligned}\tag{30}$$

It is understandable why in the expression for the polarization energy the coefficients of the terms that do not involve x are small. These terms arise out of the electronic polarization of the positive ions which have rather small polarizability. Combining Eqs. (9), (16), and (30) we get the following equation for the net change in the lattice energy, namely,

$$\begin{aligned}\Delta E_L(x,y,z)=(\alpha_M e^2/a)+(e^2/a)[&-0.538-0.374x\\ &-1.416y+2.348z+\alpha x^2+\beta y^2+\gamma z^2\\ &+\lambda xy+\mu yz+\nu xz], \end{aligned}\tag{31}$$

where the coefficients α, β, γ, λ, μ, and ν depend on our choice of the repulsive energy formula. In Table IV, the values of these coefficients are tabulated when the Born-Mayer, Huggins-Mayer, and Pauling repulsive energy formulas are used. The coefficients in the lattice-energy expression are seen to be sensitive to the choice of the repulsive energy formula, as discussed above.

From Eq. (31), one obtains the energy of the lattice as a function of x alone by minimizing the lattice energy with respect to y and z. This leads to the equations

$$\begin{aligned} 2\beta y+\mu z&=1.416-\lambda x,\\ \mu y+2\gamma z&=-2.348-\nu x. \end{aligned}\tag{32}$$

One can now solve Eqs. (32) in terms of x and substitute the following expressions (33) obtained for $y(x)$ and $z(x)$ in Eqs. (31) and hence express the lattice energy in terms of x alone, as in (34):

$$\begin{aligned} y_0(x)&=b+cx,\\ z_0(x)&=d+ex, \end{aligned}\tag{33}$$

and

$$\Delta E_L(x)=A+Bx+Cx^2. \tag{34}$$

From Eq. (34), one can obtain a plot of the lattice energy as a function of x or of R, the distance between the ions A and B, the relation between R and x being given by

$$R=\sqrt{2}a(1-\sqrt{2}x). \tag{35}$$

In Table V, the values of b, c, d, and e in Eq. (33) and A, B, and C in Eq. (34) are tabulated for our three different repulsive force laws. In Fig. 2, we have plotted $\Delta E_L(R)$ as a function of R and as expected there is great similarity between the Born-Mayer and Pauling curves, there being a minimum in the lattice energy curve in each case. The positions of the minima can be expressed either in terms of x or in terms of R, using Eq. (35). In Table VI, we have tabulated for the cases of the three repulsive formulas employed, the values of x_0, R_0, and $\Delta E_L(R_0)$, the lattice energy at the position of the minimum in $\Delta E_L(R)$. The minimum in the lattice energy curve is rather significant and arises out of a compromise between the energy of interaction between the ions A and B with the ions C, D, E, F, G, H. Thus, the distance between the ions A and C decreases as x increases while the distances between ion A and ions E and F increases as x increases. The fact that we get a minimum in the lattice energy curve indicates that there is self-trapping of the hole even in the absence of appreciable electronic conjugation between the ions A and B leading to the F_2^- molecule ion. As we shall see in the next section, there is also a stable state for the free F_2^- ion. The relative depths and shapes of the lattice energy and electronic energy curves are important in determining the nature of the total energy curve for the crystal. We shall discuss this point further in Secs. IV and V after describing the electronic energy calculation.

TABLE IV. Values of coefficients in lattice energy expressions.

Coefficient	Born-Mayer	Huggins-Mayer	Pauling
α	1.953	1.135	2.381
β	3.691	2.873	4.119
γ	8.368	6.732	9.224
λ	−0.698	−0.053	−1.678
μ	−1.707	−1.707	−1.707
ν	−2.162	−1.169	−3.288

IV. ELECTRONIC ENERGY CALCULATIONS

The next quantities that we have to calculate are the energies and wave functions for a F_2^- ion as a function of internuclear separation. There are various standard methods in molecular physics to accomplish this, broadly divided into molecular orbital and valence-

TABLE V. Coefficients b, c, d, e, A, B, and C in Eqs. (34) and (35) for the different choices of repulsive energy formula. A, B, and C are in Ry ($e^2/2a_0$).

Coefficient	Born-Mayer	Huggins-Mayer	Pauling
b	0.163	0.202	0.148
c	0.127	0.036	0.245
d	−0.124	−0.149	−0.114
e	0.142	0.091	0.201
A	0.503	0.473	0.515
B	−0.116	−0.111	−0.131
C	0.923	0.569	0.971

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VI. Positions R_0 of the minima and minimum values $\Delta E_L(x_0)$ of the lattice energy for the various approximations. $\Delta E_L(x_0)$ is expressed in Ry, and R_0 is in units of a_0. For $x=0$, that is, for the undisplaced positions of A and B, $R=5.371a_0$.

Quantity	Born-Mayer	Huggins-Mayer	Pauling
x_0	0.063	0.097	0.068
R_0	4.893	4.680	4.855
$\Delta E_L(x_0)$	0.502	0.470	0.513

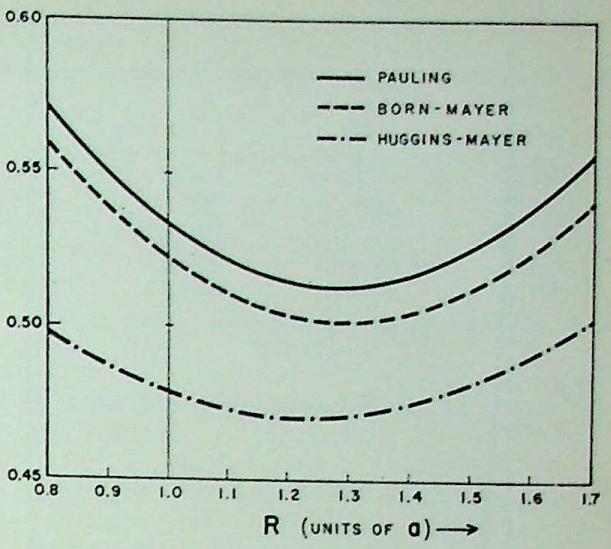


FIG. 2. Curve showing lattice energy ΔE_L as a function of the distance R between the F^- ions A and B. Curves for all three approximations for the repulsive energy, namely Born-Mayer, Pauling, and Huggins-Mayer are given.

bond methods.[14] Both these methods in their improved forms, including configuration interaction in the molecular orbital approach and ionic character in the valence bond approach, are essentially equivalent. In their improved forms they are comparable to variational approaches in the self-consistent procedure for atoms including configuration interaction. In our calculations, we have not attempted to include any configuration interaction or self-consistency requirements, but have adopted a simple molecular orbital procedure. In the usual molecular orbital method, as in the Hartree-Fock approach for atoms, one generally starts from the following many-electron Hamiltonian and computes from it a one-electron Hamiltonian for each state:

$$\mathcal{H}=-\sum_i \nabla_i^2+\sum_i \frac{2z}{r_i}+\sum_{i>j} \frac{2}{r_{ij}}. \tag{36}$$

For simplicity, we have adopted the following approximate one-electron Hamiltonian for the three $2p$ electrons which take part in the σ bond for the F_2^- molecule:

$$\mathcal{H}=-\nabla_i^2+V_A(\mathbf{r}_{Ai})+V_B(\mathbf{r}_{Bi}). \tag{37}$$

The potentials $V_A(\mathbf{r}_{Ai})$ and $V_B(\mathbf{r}_{Bi})$ both represent the one-electron potentials seen by a $2p$ electron on a F^- ion. The justification for choosing (37) is the following. When the unpaired electron in the molecule is in the neighborhood of ion A, then as far as the ion A is concerned, it sees the potential that a $2p$ electron on an F^- ion would see. In the same situation, as far as ion B is concerned, the charge distribution on it looks like a neutral atom to the unpaired electron on A and therefore the electron again experiences the same potential from ion B as a $2p$ electron on an F^- ion would. These arguments are valid when the cores of the F^- are not too distorted by the presence of the hole. This would be true if the distance between the fluorine nuclei in the molecule is not very greatly different from the distance in the perfect crystal.

As basis functions for our determination of the σ_g and σ_u functions for the F_2^- molecule, we have selected the following linear combinations:

$$\Phi_g=\sum_i c_i\phi_{ig}, \qquad \Phi_u=\sum_i d_i\phi_{iu}, \tag{38}$$

where

$$\begin{aligned} &\phi_{1g}=\phi_{1A}-\phi_{1B}, \quad \phi_{1u}=\phi_{1A}+\phi_{1B}, \\ &\phi_{2g}=\phi_{2A}-\phi_{2B}, \quad \phi_{2u}=\phi_{2A}+\phi_{2B}, \\ &\phi_{3g}=\phi_{3A}+\phi_{3B}, \quad \phi_{3u}=\phi_{3A}-\phi_{3B}, \\ &\phi_{4g}=\phi_{4A}+\phi_{4B}, \quad \phi_{4u}=\phi_{4A}-\phi_{4B}, \end{aligned} \tag{39}$$

with

$$\begin{aligned} &\phi_{1A}=\frac{P_{2p}(r_A)}{r_A}Y_1^0(\theta_A,\phi_A), \quad \phi_{3A}=\frac{P_{2p}(r_A)}{r_A^2}Y_0^0(\theta_A,\phi_A), \\ &\phi_{2A}=P_{2p}(r_A)Y_1^0(\theta_A,\phi_A), \quad \phi_{4A}=\frac{P_{2p}(r_A)}{r_A}Y_0^0(\theta_A,\phi_A). \end{aligned} \tag{40}$$

The function $P_{2p}(r_A)$ is taken to be the radial part of the $2p$ wave function of the F^- ion. Ideally one should have used in addition to ϕ_{1A} and ϕ_{2A}, functions whose radial parts are $r_AP_{2p}(r_A)$, $r_A^2P_{2p}(r_A)$, $\cdots$ to include a complete set. We have chosen only two functions out of the complete set as a compromise to reduce the work involved. The same remark applies to the choice of the functions ϕ_{3A} and ϕ_{4A}. Taking the radial parts of all the functions as being related to $P_{2p}(r_A)$ enables us to make use of the known integral and differential properties [Eq. (43)] of this function, especially as they relate to the effective potentials V_A and V_B and the atomic eigenvalues $\mathcal{E}_{2p}$.

To set up secular equations to obtain the coefficients c_i and d_i in Eq. (38), we have to compute matrix ele-

[14] C. Coulson, *Valence* (Oxford University Press, New York, 1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

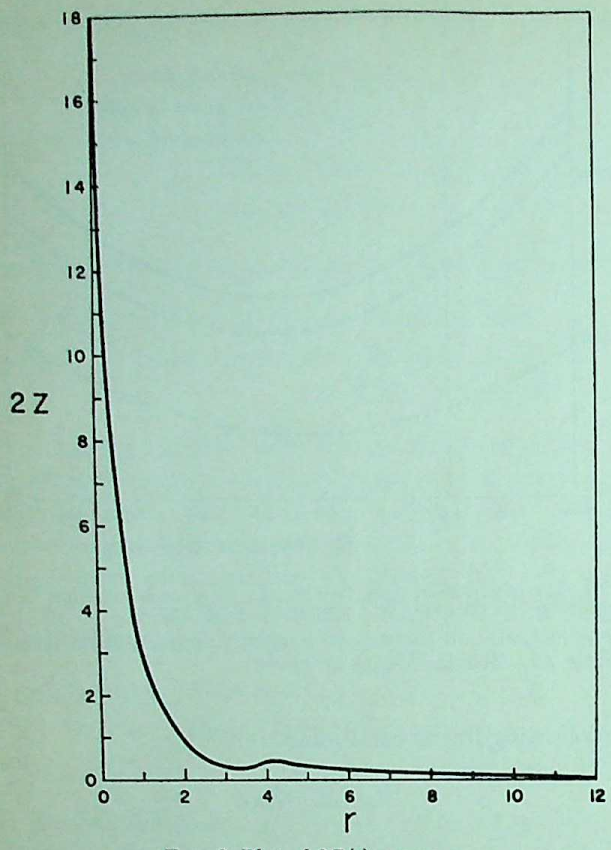


FIG. 3. Plot of $2Z(r)$ versus r.

ments $\mathcal{H}_{ij}=\langle\phi_i|\mathcal{H}|\phi_j\rangle$ for both u and g states. These matrix elements $\mathcal{H}_{ij}$ involve integrals over the orbitals in Eq. (39). Examples of the various types of integrals

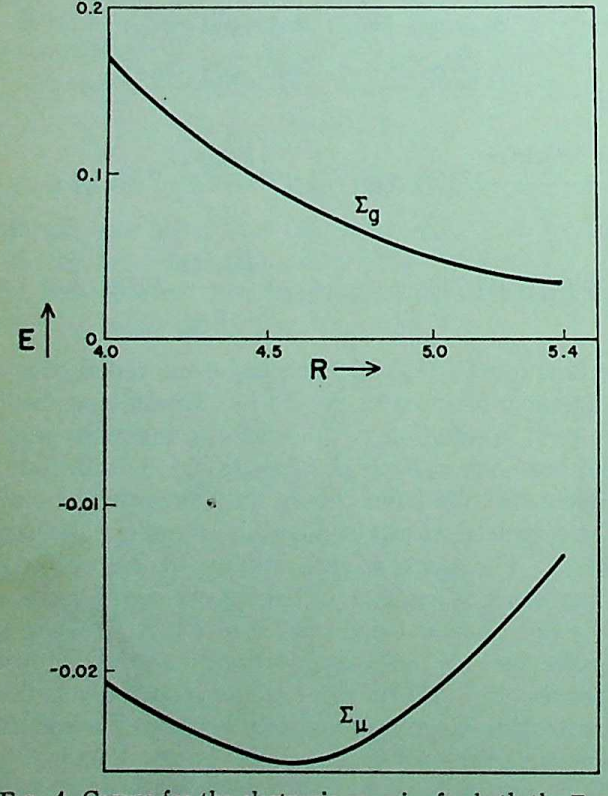


FIG. 4. Curves for the electronic energies for both the Σ_g and Σ_u states of the hole as a function of R, the distance between ions A and B.

that have to be calculated are

$$\langle\phi_{iA}|-\nabla^2|\phi_{iA}\rangle,\quad \langle\phi_{iA}|V_A|\phi_{iA}\rangle,\quad \text{and}$$
$$\langle\phi_{iA}|V_B|\phi_{iA}\rangle,\quad \langle\phi_{iB}|-\nabla^2|\phi_{iA}\rangle,\quad \langle\phi_{iA}|V_A|\phi_{iB}\rangle.$$

Of these integrals only the first two are one-center integrals, the other three are two-center integrals. The one-center integrals can be obtained by direct numerical integration. For example, we have

$$\langle\phi_{1A}|-\nabla^2|\phi_{1A}\rangle=\int_0^\infty\left(\frac{dP_{2p}}{dr}\right)^2dr+2\int_0^\infty\frac{P_{2p}^2}{r^2}dr,$$
$$\langle\phi_{1A}|V_A|\phi_{1A}\rangle=-\int_0^\infty\frac{P_{2p}^2}{r}2Z(r)dr, \tag{41}$$

where the potential $V_A(r)$ has been expressed as

$$V_A(r)\equiv-2Z(r)/r. \tag{42}$$

We obtain $V_A(r)$ directly from its definition by the one-electron Hartree-Fock equation:

$$V_A(r)=+\mathcal{E}_{2p}-\frac{P_{2p}''}{P_{2p}}-\frac{2}{r}, \tag{43}$$

where $\mathcal{E}_{2p}$ is the one-electron energy for the $2p$ state and $P''=d^2P/dr^2$. The function P_{2p} and the energy parameter $\mathcal{E}_{2p}$ are taken from Froese's calculations.[8] In Fig. 3 we have plotted the values of $2Z(r)$ as a function of r. For the two-center integrals we employ Löwdin's α-function method.[15] The α functions and the integrals involving them that were required [to determine the matrix elements of the Hamiltonian $\mathcal{H}$, Eq. (37)], were obtained at the computing centers of the University of Rochester and Argonne National Laboratory. Since we require the electronic energy as a function of the distance between the ions A and B we have to compute the matrix elements for different values of R. These computations were done for six values of R, namely $4.0a_0$, $4.4a_0$, $4.8a_0$, $5.0a_0$, $5.2a_0$, and $5.4a_0$. Since the functions ϕ_i do not form an orthonormal set, we must take account of the integrals $S_{ij}=\langle\phi_i|\phi_j\rangle$ for both the σ_g and σ_u states and the eigenvalue secular equation then takes the form:

$$\mathrm{Det}|\mathcal{H}_{ij}-S_{ij}E|=0 \tag{44}$$

for both σ_g and σ_u states. The eigenvalues $\mathcal{E}_g$ and $\mathcal{E}_u$ for the six values of R considered are tabulated in Table VII. Also tabulated in Table VII are the quantities

$$E_u=\mathcal{E}_u+2\mathcal{E}_g+(2/R)-3\mathcal{E}_{2p} \tag{45}$$

and

$$E_g=2\mathcal{E}_u+\mathcal{E}_g+(2/R)-3\mathcal{E}_{2p}. \tag{46}$$

The quantity E_u represents the decrease in energy of the three σ electrons due to bonding between the ions

[15] P. O. Löwdin, Advan. Phys. 5, 96 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VII. Energy values in rydbergs for electronic and hole states for different values of the internuclear distance of R.

Energy \ R value	4.0	4.4	4.8	5.0	5.2	5.4
ε_g	−0.5997	−0.5666	−0.5400	−0.5282	−0.5179	−0.5086
ε_u	−0.4103	−0.4352	−0.4498	−0.4534	−0.4558	−0.4552
$\varepsilon_u+2\varepsilon_g+(2/R)$	−1.1097	−1.1137	−1.1131	−1.1098	−1.1070	−1.1020
$\varepsilon_g+2\varepsilon_u+(2/R)$	−0.9203	−0.9823	−1.0229	−1.0350	−1.0414	−1.0486
E_u	−0.0207	−0.0247	−0.0241	−0.0208	−0.0180	−0.0130
E_g	+0.1687	+0.1067	+0.0661	+0.0540	+0.0376	+0.0304

A and B, when the hole is in the Σ_u state; E_g represents the corresponding energy decrease of the three σ electrons for the Σ_g state of the hole. The term $2/R$ represents the Coulomb energy of repulsion between the effective unit positive charges at the A and B ions about which the three σ electrons of the F_2^- molecule can be considered as moving. Thus E_u and E_g give the values of the term E_E in Eq. (1) for the Σ_u and Σ_g states of the hole, respectively. The variations of E_u and E_g with R are shown in Fig. 4. It is seen that while E_u has a minimum at $R=4.56a_0$, E_g has no such minimum. This result indicates that while the Σ_u state for the hole in the free F_2^- molecule is stable, the Σ_g state is not, as expected. Finally, to obtain the variation in total energies of the hole states Σ_u and Σ_g in the crystal with the distance R between A and B, we have to combine the energies plotted in Figs. 2 and 4 which will give us the variation with R of the total energy of the crystal in the presence of the hole.

V. TOTAL ENERGY OF CRYSTAL AND INTERPRETATION OF VARIOUS EXPERIMENTALLY STUDIED PROPERTIES OF THE HOLE

In Table VIII, we tabulate $E-E_0$ of Eq. (1) as a function of R. In addition to the lattice energy calculated from Eq. (34) and the electronic energy tabulated in Table VIII we have added the electronic affinity E_A which we have taken as 0.363 Ry from Froese's calculations.[8] For the lattice energy, in view of the inaccuracy in principle in the use of the Huggins-Mayer expression, as mentioned earlier, we have only used the Born-Mayer and Pauling expressions. Also tabulated is δE which is the difference between E and the quantity

$$\frac{\alpha_m e^2}{a}+E_A=1.283\ \text{Ry}, \tag{47}$$

representing the increase in energy of a crystal when an electron leaves the crystal (or a hole is produced in the crystal) excluding the energy change due to relaxation of the ions of the lattice and the electronic binding between the ions A and B. The suffixes g and u to δE and E are added in Table VIII to indicate the corresponding states Σ_g and Σ_u of the hole. The negative values of δE_u and δE_g indicate that the combined effects of the lattice distortion and electronic conjugation between A and B ions is to stabilize the hole center. In the case of the Σ_u state both the lattice distortion and electronic conjugation lead to stabilization. For the Σ_g state, the lattice distortion leads to stabilization while the presence of two antibonding σ_u electrons and only one bonding σ_g electron leads to destabilization. In Fig. 5 the energies δE_u and δE_g are plotted as functions of R for both the Pauling and Born-Mayer approximations. The zero ordinate on this plot would represent the energy of the undisturbed crystal after the hole has been removed, while the energy of the perfect crystal E_0 lies at −1.283 Ry. As was found in Fig. 4 with the electronic energy, again we have a minimum only in the Σ_u state and none in the Σ_g state suggesting that only the Σ_u state for the hole has a stable minimum. This result cannot be taken too literally, however, because we must realize that when R corresponds to the normal lattice distance or greater distances, our de-

TABLE VIII. Energy in Ry of the crystal due to the presence of the hole.

Energy \ R	4.0	4.4	4.8	5.0	5.2	5.4
E_u (Pauling)	0.869	0.855	0.852	0.855	0.860	0.868
E_u (Born-Mayer)	0.858	0.845	0.841	0.844	0.848	0.856
E_g (Pauling)	1.058	0.987	0.942	0.930	0.915	0.911
E_g (Born-Mayer)	1.047	0.976	0.931	0.919	0.904	0.899
δE_u (Pauling)	−0.415	−0.428	−0.431	−0.428	−0.424	−0.416
δE_u (Born-Mayer)	−0.425	−0.439	−0.442	−0.439	−0.435	−0.428
δE_g (Pauling)	−0.225	−0.297	−0.341	−0.353	−0.368	−0.372
δE_g (Born-Mayer)	−0.236	−0.307	−0.352	−0.364	−0.380	−0.384

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

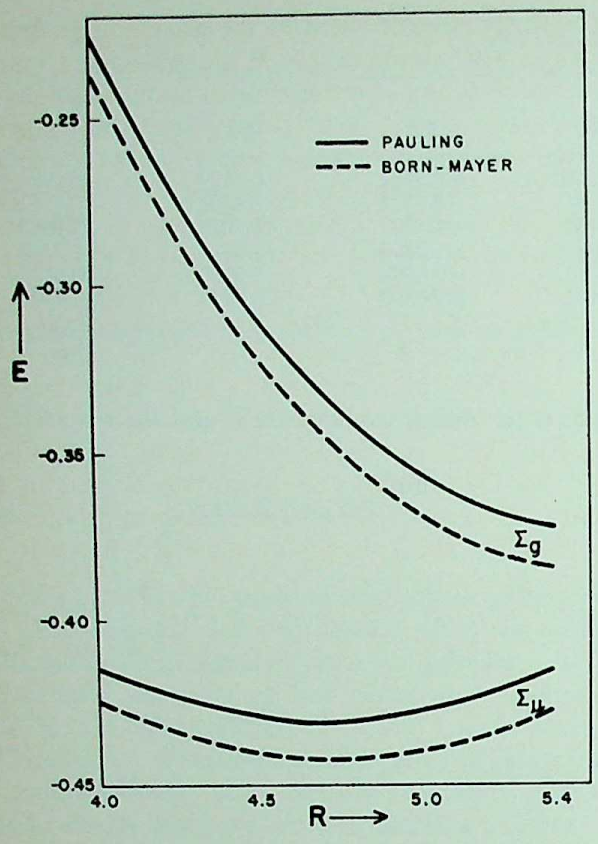


FIG. 5. Curves for the total energies δE_u and δE_g for the Σ_u and Σ_g states, respectively, of the hole as a function of R.

scription is not a complete one. In that case, there is high probability of the hole hopping to a new pair of F^- ions.

A comparison of Figs. 2, 4, and 5 indicates that the depths and rates of variation around the minimum of the Born-Mayer and Pauling curves are comparable. Also it is seen that the minima of the lattice energy curves for the Pauling and Born-Mayer approximations are at $R=4.89a_0$ and $4.92a_0$, respectively, while the minimum in the E_u curve for the electronic energy occurs at about $4.56a_0$. In comparison, the minimum in the δE_u curves for the total energy occurs at about $4.72a_0$ for both the Pauling and Born-Mayer approximations. These considerations indicate that according to our calculations, the lattice distortion around the V_k center and the electronic binding within the F_2^- molecule are comparable in importance in contributing to the stabilization of the V_k center.

From Fig. 4 one can also obtain the absorption energy for transitions between the Σ_u and Σ_g hole states. By the Franck-Condon principle, we expect the transition energy to be obtained vertically from the minimum in the E_u curve. From Fig. 5, the absorption energy for both the Born-Mayer and Pauling approximations is 0.097 Ry, which corresponds to 1.32 eV. This is to be compared with the experimentally observed[16] position of the absorption peak at 3.54 eV. We thus obtain only order-of-magnitude agreement with experiment. There are two possible sources of error which might be responsible for the disagreement between theoretical and experimental absorption energies. Our molecular calculations are approximate and perhaps a better calculation would lead to a stronger binding that is, a deeper energy curve for the Σ_u state of the hole and move the Σ_g state correspondingly upwards in energy.[14] The other possibility is the lack of self-consistency in our calculation which is implied by the direct addition of E_E and ΔE_L.

From the ground state and excited state curves for the energy as a function of R, we can also obtain the width of the absorption band. Assuming a linear dependence on the excited state configuration coordinate (R in our case)[17] one obtains the following expression for the width at an intensity equal to one half of that at the maximum, namely:

$$\Delta W = 27A\,(\mu K)^{-1/4}\ \text{eV}, \tag{48}$$

where A is the excited state slope in Ry/a_0, μ is the effective mass of the oscillator in units of the electron mass, and K is the ground-state force constant in Ry/a_0^2. Properly, one should use not only the reduced mass of the two F atoms A and B but include a small correction from the ions $C, D, \cdots, H$ which relax. However this last correction is rather difficult to calculate and is not warranted because of the various approximations that we have made so far in our calculations. So we take

$$\mu = 9.5m_e \times 1837,$$

where the electron mass, $m_e = 9.11\times10^{-28}$ g. From the Born-Mayer curve in Fig. 5, we get $K=0.1\ \text{Ry}/a_0^2$, $A=0.1\ \text{Ry}/a_0$. Substituting these numbers in (48) we have

$$\Delta W = 0.58\ \text{eV} \tag{49}$$

as compared to the experimental value of 0.95 eV found by Delbecq *et al.*[16] So the calculated value of the half-width is about a factor of 2 less than experiment. An improved electronic energy calculation would lead to a deeper potential well for the Σ_u state implying an increase in K, while the Σ_g state would get steeper leading to an increase in A. These two consequences are opposite in their influence on ΔW but perhaps their combined effect would lead to better agreement with experiment. Another quantity that can be derived from Fig. 5 from the ground-state curve is the half-width of the lowest vibrational state. One can deduce the following expression for this latter half-width, namely

$$(\Delta R)_{1/2} = (0.996)\,(\mu K)^{-1/4} a_0. \tag{50}$$

Using the values of μ and K for the Born-Mayer curve

[16] C. J. Delbecq, W. Hayes, and P. H. Yuster, Phys. Rev. 121, 1043 (1961).

[17] See, for example, D. L. Dexter in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1958), Vol. 6, p. 353.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

we get $(\Delta R)_{1/2}=0.015a_0$. The Pauling curves in Fig. 5 yield almost exactly the same values for ΔW and $(\Delta R)_{1/2}$ as the Born-Mayer curves. This small value of $(\Delta R)_{1/2}$ assures us that the actual inter-nuclear distance in the F_2^- molecule is always well within the region in which our calculation is valid.

With the wave functions that we have obtained for the V_k center, we can make an approximate calculation of the hyperfine coupling constants between the unpaired electron spin and the F^{19} nucleus. The Hamiltonian describing the magnetic hyperfine interaction between a nucleus and the spins of the orbital electrons is given by

$$\mathcal{H}=\mathcal{H}_{\rm dip}+\mathcal{H}_F, \tag{51}$$

where

$$\mathcal{H}_{\rm dip}=\frac{2\mu_N\mu_B}{I}\sum_i\left[\frac{\mathbf{S}_i\cdot\mathbf{I}}{r_i^3}+\frac{3(\mathbf{S}_i\cdot\mathbf{r}_i)(\mathbf{I}\cdot\mathbf{r}_i)}{r_i^5}\right] \tag{52a}$$

and

$$\mathcal{H}_F=(16\pi/3)(\mu_N\mu_B/I)\sum_i\mathbf{S}_i\cdot\mathbf{I}\delta(\mathbf{r}_i). \tag{52b}$$

In Eqs. (52), μ_N and $\mathbf{I}$ are, respectively, the magnetic moment and spin of the nucleus, μ_B is the Bohr magneton, $\mathbf{S}_i$ and $\mathbf{r}_i$ are, respectively, the spin and position vectors of the ith electron. The summation in i runs over all the electron spins in the atomic or molecular system under consideration. When we have a system with one unpaired electron and the wave functions for the various electronic states of the system are orthogonal we can drop the summation over i in Eqs. (52a) and (52b) and take expectation values over the orbital corresponding to the single unpaired electron. However, when the wave function for the orbital of the unpaired electron is not orthogonal to the wave functions for other states one has to take the expectation values of the summations in (52a) and (52b) over the entire determinantal wave function involving both the outer unpaired electronic state and all other paired states of the molecule. It can be shown[18] however, that in such cases, one can simplify the calculation by making the unpaired electron wave function orthogonal to the wave functions for the paired states by the Schmidt orthogonalization procedure. One then has to take the expectation value of only one term out of the summations in (52a) and (52b) over the orthogonalized wave function. For the ground state Σ_u of the hole, the unpaired electron will be in a σ_u state. The orthogonalized wave function for the unpaired electron will therefore be of the form:

$$\psi=\frac{\psi_u-S_1(\psi_{1SA}+\psi_{1SB})-S_2(\psi_{2SA}+\psi_{2SB})}{N[1+2S_1^2\langle\psi_{1SA}|\psi_{1SB}\rangle+2S_2^2\langle\psi_{2SA}|\psi_{2SB}\rangle+4S_1S_2\langle\psi_{1SA}|\psi_{2SB}\rangle]^{1/2}}, \tag{53}$$

where

$$S_1=\langle\psi_u|\psi_{1SA}\rangle=-\langle\psi_u|\psi_{1SB}\rangle,$$
$$S_2=\langle\psi_u|\psi_{2SA}\rangle=-\langle\psi_u|\psi_{2SB}\rangle,$$
$$N^2=\langle\psi_u|\psi_u\rangle.$$

The overlap integrals $\langle\psi_{1SA}|\psi_{1SB}\rangle$, $\langle\psi_{1SA}|\psi_{2SB}\rangle$, and $\langle\psi_{2SA}|\psi_{2SB}\rangle$ are all one order of magnitude less than the overlap integrals S_1 and S_2, so we shall neglect them. Using Eq. (53) for the wave function ψ of the unpaired electron and Eq. (52a) for $\mathcal{H}_{\rm dip}$, we obtain Eq. (54) for the hyperfine constant "b". Since we want to express "b" in Oersteds we have divided the hyperfine constant in energy units (ergs) by $2\mu_B$.

$$\begin{aligned}b=(3\mu_N/2IN^2)[&\langle\psi_u|0|\psi_u\rangle+S_1^2\langle\psi_{1SA}|0|\psi_{1SA}\rangle\\&+S_2^2\langle\psi_{2SA}|0|\psi_{2SA}\rangle-2S_1\langle\psi_u|0|\psi_{1SA}\rangle\\&-2S_2\langle\psi_u|0|\psi_{2SA}\rangle+2S_1S_2\langle\psi_{1SA}|0|\psi_{2SA}\rangle],\end{aligned} \tag{54}$$

with

$$0=(3\cos^2\theta_A-1)/r_A^3,$$

where r_A and θ_A refer to the polar coordinates of the electron with respect to A and with the AB direction as the Z axis. Equation (54) is approximate because a number of cross terms, involving $1S$ and $2S$ orbitals on atom B have been dropped. These terms are three orders of magnitude less than the leading term $\langle\psi_u|0|\psi_u\rangle$. Furthermore it is clear that the second, third, and sixth terms in (54) are identically zero. The fourth and fifth terms are two orders of magnitude smaller than the leading term. Therefore we can write, using (38) and (39),

$$\begin{aligned}b=(3\mu_N/2IN^2)[&\sum_{ij}d_id_j\{\langle\phi_{iA}|0|\phi_{jA}\rangle\\&+\langle\phi_{iA}|0|\phi_{jB}\rangle+\langle\phi_{iB}|0|\phi_{jA}\rangle+\langle\phi_{iB}|0|\phi_{jB}\rangle\}].\end{aligned} \tag{55}$$

In Eq. (55) the fourth term in brackets is two orders of magnitude smaller than the contribution from the first term while the second and third terms are each more than one order of magnitude smaller. Also, when either i or j is three or four, the contribution to the first term is identically zero. We thus have finally,

$$\begin{aligned}b=(6\mu_N/5IN^2)[&d_1^2\langle\phi_{1A}|r_A^{-3}|\phi_{1A}\rangle+d_2^2\langle\phi_{2A}|r_A^{-3}|\phi_{2A}\rangle\\&+2d_1d_2\langle\phi_{1A}|r_A^{-3}|\phi_{2A}\rangle].\end{aligned} \tag{56}$$

Using Eq. (39) for ϕ_{1A} and ϕ_{2A} we then obtain for b the value 747.3 Oe for both the Pauling and Born-Mayer approximations. More than 99% of the contribution to b comes from the first term in (56). Since our wave functions were tabulated for the points listed in Table VIII, we had to interpolate the coefficients d_i in Eq. (55) at R equal to $4.72a_0$. The calculated value

[18] See, for example, B. Gourary and F. J. Adrian, Phys. Rev. 105, 1180 (1957); W. Blumberg and T. P. Das, *ibid.* 110, 647 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IX. Contributions from the various terms in Eq. (57) to the isotropic hyperfine interaction constant for F^{19} nucleus in the V_k center. The wave function densities are in units of a_0^{-3} and the contributions to a_F in Oe.

Term in Eq. (57)	Wave function density	Contribution to hyperfine constant a_F
First	0.00012	0.175
Second	0.00684	10.317
Third	0.00604	9.099
Fourth	−0.00178	−2.679
Fifth	−0.00167	−2.516
Sixth	0.01288	19.384
Total	0.02243	33.780

of the hyperfine constant b is in fair agreement, certainly in much better agreement, with experiment than the optical absorption frequency and width. The experimental value of b is not known unambiguously[3] because it depends on the sign of a which is not known. If the sign of a is assumed negative, the experimental value of b would be 946 Oe while for a positive the value would be 828 Oe. The hyperfine constant a is given by

$$a=a_F-(b/3), \tag{57}$$

where

$$a_F=(8\pi\mu_N/3IN^2)[\psi_u^2(A)+S_1^2\psi_{1SA}^2(A)+S_2^2\psi_{2SA}^2(A)-2S_1\psi_u(A)\psi_{1S}(A)-2S_2\psi_u(A)\psi_{2S}(A)+2S_1S_2\psi_{1S}(A)\psi_{2S}(A)].$$

The term a_F arises from the Fermi contact term $\mathcal{H}_F$ in the Hamiltonian and $-b/3$ from $\mathcal{H}_{\rm dip}$. In Eq. (57) we have neglected all the terms that involve $\psi_{1SB}(A)$ and $\psi_{2SB}(A)$. We have done this because the terms involving $\psi_{1SB}(A)$ and $\psi_{2SB}(A)$ are an order of magnitude smaller than the smallest of the terms in Eq. (54). This approximation is in line with our assumption concerning the smallness of the overlaps between the $1S$ and $2S$ orbitals about the two centers. To emphasize the importance of the overlap terms which may also be referred to as the Pauli correlation terms,[19] we have tabulated in Table IX the values of the various terms in the brackets of Eq. (57) in units of a_0^{-3} as well as the value of a_F in Oersteds for $R=4.72a_0$, the equilibrium distances for both the Pauling and Born-Mayer cases.

Combining a_F and $-b/3$ we get for a the value -215.3 Oe. It appears that the constant b is in much better agreement with experiment than the costant a. The reason for this can be gauged by examining Table IX. From Table IX it appears that the major contribution to a comes from the second, third, and sixth terms in Eq. (57). All these terms depend bilinearly on the overlap integrals S_1 and S_2. A stronger conjugation between the two ions A and B, as expected from earlier considerations of the optical data, would reduce the distance corresponding to the minimum in the total energy curve in Fig. 4 and would therefore increase S_1 and S_2 significantly. Such an increase would enhance a_F and therefore cause a greater cancellation between a_F and $-b/3$ leading to a smaller theoretical value of a. The value of b on the other hand, as seen from Eq. (56) does not depend sensitively on R and should therefore be relatively unaffected by a stronger conjugation. This stronger conjugation would also make the electronic energy curve deeper than the lattice energy curve and therefore make the properties of the V_k center less sensitively independent on the lattice potential. Evidence that this is the case has been obtained by Twidell and Hayes[20] in recent work on CaF_2. It should be mentioned here that in addition to the Pauli correlation effect there will be a contribution to a_F from the exchange polarization effect[21] which arises out of the difference in the potentials experienced by electrons in the two spin states of the $1S$ and $2S$ shells. For an order of magnitude estimate of the exchange polarization one could use as a criterion the results from the nitrogen atom. Thus, noting that there are three unpaired $2p$ electrons in the nitrogen atom as compared to approximately one-half of a $2p$ electron per fluorine atom in the present case, one would expect the exchange polarization contribution to a to be about a sixth of the 10 Mc result found for the nitrogen atom. This would correspond to only about 0.25 Oe. While this result is probably an underestimate because the $1S$ and $2S$ wave functions in fluorine atom are more strongly bound than in nitrogen, it is however clear that a consideration of exchange polarization effect alone cannot explain the difference between our calculated values of a and experiment.

VI. CONCLUSION

From the calculations in the previous sections it appears that both the distortion of the lattice around the V_k center and the electronic binding of the σ electrons play important roles in determining the properties of the V_k center. The relative importance of the lattice and electronic energies would of course determine how the properties of the V_k center will vary from one crystal to another. Recent work of Twidell and Hayes,[20] however, indicates that the electronic properties of the V_k center in CaF_2 are almost the same as in LiF. This could be explained either if the electronic binding energy of the F_2^- molecule and the lattice energy are as comparable with each other in importance as in LiF or if the electronic binding energy has been underestimated by us and really involves a curve with a steeper minimum than we have calculated and the lattice effect is of minor importance in determining the equilibrium

[19] A. Mukherjee and T. P. Das, Phys. Rev. 111, 1479 (1958).

[20] A. Twidell and W. Hayes, Proc. Phys. Soc. (London) A79, 1295 (1962).

[21] V. Heine, Phys. Rev. 107, 1002 (1957); T. P. Das and A. Mukherjee, J. Chem. Phys. 33, 1808 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

configuration. The latter possibility seems somewhat more probable, from our considerations of optical and hyperfine data in the last section. Also, we have made an approximate calculation of the energy of a trapped fluorine atom at a F^- site using the Born-Mayer approximation and allowing for displacements of the nearest neighbor Li^+ ions. The procedure employed in the calculations is analogous to that used by Dick[22] in recent work on alkali halide solid solutions. The energy of the trapped fluorine atom relative to the undisturbed crystal after the hole is removed is then found to be -0.441 Ry. Comparing this with the corresponding energy, -0.443 Ry for the V_k center for the Born-Mayer approximation, it is noticed that they are almost equal, although one would expect the V_k center to be lower in energy from a consideration of relative stability. This result seems to substantiate our conclusion that we have underestimated the electronic binding energy of the V_k center and should have a curve with a steeper minimum for the ground state in Figs. 4 and 5.

It would be useful to combine the results of a careful recalculation[23] of the electronic energy of the F_2^- molecule with the results of our lattice energy calculations in Sec. III to obtain a better total energy curve for the V_k center. It would also be interesting to find out if the recalculated wave function for the electronic σ_u state gives better agreement with the experimental value of the hyperfine constant a. It is encouraging that the results of the lattice energy calculations in Sec. III with two widely different repulsive energy formulas, such as the Born-Mayer and Pauling expressions (10) and (12), are not very different from one another and especially that the curves in Figs. 2 and 5 are so similar in shape. This is a very significant result because the effects of such widely different force laws have not been extensively compared in imperfect crystals. In future calculations one could therefore use either the Born-Mayer or Pauling repulsive formula with equal confidence.

One possible source of error in the calculation of the absorption energy should be mentioned here. When we consider the Hartree-Fock approximation for an atomic or molecular system, the net energy of the system is not just the sum of the one-electron energies for the individual states because then the Coulomb and exchange interaction energies get counted twice. One has therefore to subtract off the Coulomb and exchange energies from the sum of one-electron energies for the various states. The difference in E_u and E_g, which are tabulated in the last column of Table VII, is therefore likely to differ from the $\Sigma_u \rightarrow \Sigma_g$ excitation energy for the hole if there is appreciable difference in the exchange and Coulomb integrals occurring in the energy expressions for the σ_g and σ_u states of the electrons. We have not made any estimates of the correction due to this source of error because such a refinement would be inconsistent with the approximations we have already made in our electronic energy calculations in Sec. IV.

There is one other correction that has to be considered in any improved calculations in future. In Sec. IV and in Fig. 5 we have directly added the electronic and lattice energies associated with the V_k center which amounts to a neglect of self-consistency requirements. To be entirely self-consistent, one should consider the change in the potential seen by the electrons due to the distortion and polarization of the lattice. Correspondingly, one has also to consider the change in lattice energy produced by the redistribution of the electron density on the F_2^- molecule due to electronic binding. However, if it turns out in a refined calculation that the electronic binding is much more important than the lattice distortion in determining the equilibrium situation for the V_k center, one would expect the consistency condition involving interaction between lattice and electronic potentials to be correspondingly less important.

ACKNOWLEDGMENTS

This work was initiated when two of the authors (T. P. Das and R. S. Knox) were at the University of Illinois. We are grateful to Professor F. Seitz for his hospitality and his interest in this work. We are also grateful to Professor B. G. Dick of the University of Utah and Dr. T. Gilbert and many other colleagues at Argonne National Laboratory and the University of Rochester for valuable discussions. Dr. Yamashita kindly communicated to us his earlier unpublished work in this field which was rather helpful to us.

[22] B. G. Dick and T. P. Das, Phys. Rev. 127, 1053 (1962).

[23] See, for example, A Kolos and C. J. Roothaan, Rev. Mod. Phys. 32, 219 (1960); B. Ransil, Rev. Mod. Phys. 32, 245 (1962). Such calculations are already in progress at Argonne National Laboratory (A. C. Wahl and T. L. Gilbert, private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Electron Tunneling Through Thin Aluminum Oxide Films*

THOMAS E. HARTMAN AND JAY S. CHIVIAN
Texas Instruments Incorporated, Dallas, Texas
(Received 13 December 1963)

The thin-film Al-Al_2O_3-Al system is studied. The current density as a function of temperature and voltage through aluminum oxide less than 30 Å thick is analyzed using the theory of Stratton. The charge transport mechanism is established as tunneling by the functional form of the small temperature dependence. By fitting the current density data to the temperature dependence for $79°K < T < 300°K$ and $0 \leq V \leq 1.0$ V, the integrals arising from the WKB approximation are evaluated for an arbitrary barrier shape. The region of applied voltage for which quadratic power series expansions in V of these integrals are valid is established as $0.0 \leq V \leq 0.3$ V. Assuming a trapezoidal barrier shape, approximate values for the barrier heights of 0.89 and 0.78 eV and an electron effective mass of $m^*/m = 1$ are deduced. Comparisons are drawn between the more general forms of Stratton's and Simmons' theories.

I. INTRODUCTION

THE flow of electrons through a sufficiently thin insulating film will occur by the quantum mechanical mechanism of tunneling.[1] Barrier layers thicker than a few tens of angstroms are nearly impenetrable to tunneling electrons, as the transmission probability falls off exponentially with thickness, except where the tunnel distance is shortened by the application of a sufficiently large voltage (i.e., Fowler-Nordheim[2,3] region). For diminishingly small applied voltages, the electronic conductivity of films a few tens of angstroms thick may arise from thermionic or Schottky emission over the barrier,[4] tunneling through the barrier, space-charge limited current,[5] or a combination of these mechanisms with traps in the forbidden gap of the insulator participating as an intermediary.[6]

At temperautres below the superconducting transition for the aluminum films in the Al-Al_2O_3-Al system, the electron current through an aluminum oxide film a few tens of angstroms thick has been shown to result from tunneling.[7] The thermionic or Schottky contribution to the current is clearly negligible at these temperatures.

Emtage and Tantraporn[4] have reported that for the Al-Al_2O_3-Al system the entire current through a 50-Å-thick oxide film at room temperature appears to come from Schottky emission. Whether thermionic emission or tunneling dominates the zero-voltage conductivity depends very strongly on the thickness of the oxide film. For sufficiently thin oxide films tunneling will be the dominant conduction mechanism. It is easy to show for the Al-Al_2O_3-Al system, assuming a rectangular barrier 0.85 eV high and an effective mass $m^*/m = 1$, that at room temperature the critical thickness is 35 Å (or thinner), tunneling will dominate almost exclusively. For oxide films thicker than 35 Å, thermionic or Schottky emision will dominate and the conductivity will be highly temperature-dependent.

The current density through aluminum oxide films less than 30 Å thick, sandwiched between two aluminum films, was measured as a function of temperature and voltage. The conduction mechanism, even for diminishingly small applied voltages, is shown to be tunneling by the functional form of relatively small temperature dependence. The data are analyzed using the theory of Stratton.[8] Using the barrier height and thickness deduced from the data the Schottky current was calculated and found to be more than two orders of magnitude less than the measured current at room temperature. The integrals arising from the WKB approximation in the theory are evaluated for an arbitrary potential barrier shape by fitting the current density data to the temperature dependence. Finally, the theory of Stratton is modified to take more accurate account of the back current, which is the current that flows in the direction opposite to the current which is aided by the applied potential. The resulting equation can be reduced to either Stratton's or Simmons' theory depending on the approximations employed.

II. EXPERIMENTAL TECHNIQUE

A. Sample Preparation

The configuration of the aluminum-aluminum oxide-aluminum thin-film sandwiches employed in these experiments is shown schematically in Fig. 1. The substrate is a glass microscope slide cleaned by a technique similar to that of Lowe.[9] The slide is subsequently ion bombarded in the manner described by Holland.[10]

The metal layers are deposited on an unheated substrate in a vacuum-evaporation system at a pressure of less than 1×10^{-6} Torr. The composition of the resi-

* Presented in part at the Pasadena meeting of the American Physical Society, December 1963.

[1] R. Holm, *Electric Contacts* (Hugo Gerer, Stockholm, Sweden, 1946).

[2] R. Fowler and L. Nordheim, Proc. Roy. Soc. (London) A119, 173 (1928).

[3] C. A. Mead, J. Appl. Phys. 32, 646 (1961).

[4] P. R. Emtage and W. Tantraporn, Phys. Rev. Letters 8, 267 (1962).

[5] D. V. Geppert, J. Appl. Phys. 33, 2993 (1962).

[6] C. A. Mead, Phys. Rev. 128, 2088 (1962).

[7] I. Giaever and K. Megerle, Phys. Rev. 122, 1101 (1961).

[8] R. Stratton, Phys. Chem. Solids 23, 1177 (1962).

[9] L. Holland, *Vacuum Deposition of Thin Films* (J. Wiley & Sons Inc., New York, 1958), p. 71.

[10] L. Holland, *Vacuum Deposition of Thin Films* (J. Wiley & Sons Inc., New York, 1958), p. 84.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Residual gas analysis.[a]

Gas or vapor	Partial pressure, % of total pressure
Water vapor	89
Carbon monoxide	8
Nitrogen	1
Oxygen	1
dc 704 (pump oil)	<1

[a] The residual gas analysis was made with a mass-spectrometer similar to the one described by W. D. Davis and T. A. Vanderslice, *Vacuum Symposium Transactions*, edited by C. R. Meissner (Pergamon Press, Inc., London, 1960), p. 417.

dual gasses in the system at its ultimate pressure of 3×10^{-8} Torr are shown in Table I. The silver used for contacts and interferometry gauges is vaporized from a molybdenum Knudsen surface emitting source[11] and the 99.999+% pure aluminum is vaporized from a tungsten hairpin filament source.[12]

First the lower aluminum layer is deposited on aluminum contacts. The slide is then removed from the vacuum system, and placed in a preheated furnace of dry air at atmospheric pressure and a temperature of 300°C for two to four hours.[13] The slide is then vapor degreased in isopropyl alcohol, replaced in the vacuum system, ion bombarded, and the remaining metal layers are deposited.

The metal films are each approximately 1000 Å thick. The junction area is 10^{-2} cm^2, and is coated with a protective layer of silicon monoxide 2000 to 3000 Å thick (not shown in Fig. 1) using a tantalum perforated chimney source similar to the one described by Drumheller.[14] The samples are stored in a vacuum desiccator at room temperature.

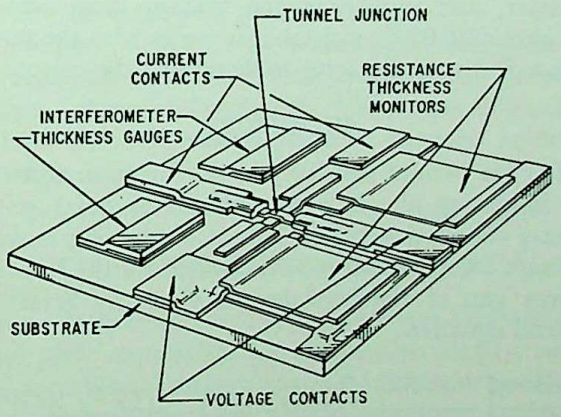


FIG. 1. Sample configuration.

[11] L. Holland, *Vacuum Deposition of Thin Films* (John Wiley & Sons, Inc., New York, 1958), p. 119.

[12] S. Dushman, *Scientific Foundations of Vacuum Technique*, edited by J. M. Lafferty (John Wiley & Sons, Inc., New York, 1962), p. 706.

[13] The rate of oxidation of evaporated aluminum films in very dry atmospheres of air or oxygen is observed to be much slower than when water vapor is present in small quantities in agreement with the results of P. E. Blackburn and E. A. Gulbransen, J. Electrochem. Soc. 107, 944 (1960) and I. Giaever and K. Megerle, Phys. Rev. 122, 1101 (1961).

[14] C. E. Drumheller, *Vacuum Symposium Transactions*, edited by L. E. Preuss (Pergamon Press, Inc., London, 1961), p. 306.

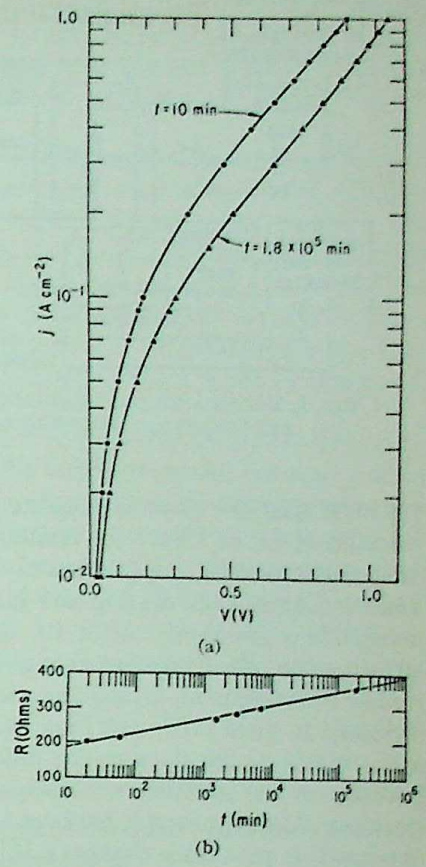


FIG. 2. Electrical characteristics of a typical sample as a function of time. (a) Initial and final plot of the current density as a function of the applied voltage. (b) Plot of the zero bias junction resistance as a function of time.

B. Sample Characteristics

For applied voltages less than 90 mV, the electrical characteristics of the samples are Ohmic (see Fig. 7). The resistance and capacitance of the samples are measured using a low-frequency-transformer ratio arm bridge which impresses 90 mV rms (or less) upon the sample, hence the values obtained correspond essentially to the zero-applied-voltage condition. All other electrical measurements are made using four-terminal methods, so that lead and contact resistance errors are avoided. The resistance of the samples range from 100 to 400 ohms and their capacitance is about 0.022 to 0.028 μF. The zero-applied-voltage resistance R and the j versus V characteristics for a typical sample are shown in Fig. 2, where j is the current density, V is the applied voltage, and t is the time ($t=0$ is the time when the sample fabrication was completed).

An increase in the junction resistance of the same general form as that reported by Handy[15] is observed. As shown in Fig. 2(b) the junction resistance increases approximately as the logarithm of the time. Extrapolating this plot the junction resistance would be expected to be double its initial (i.e., $t=20$ min) value at $t=1.6\times10^6$ min (or in about three years). The slope of the junction resistance versus $\log t$ is similar to that reported by Handy for unprotected junctions within 2 to 3 min after deposition of the second aluminum layer. Handy

[15] R. M. Handy, Phys. Rev. 126, 1968 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

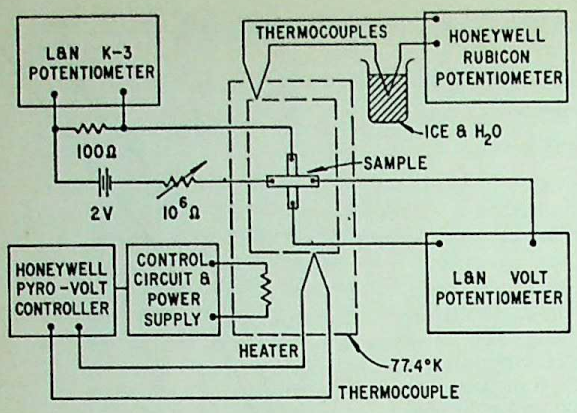


FIG. 3. Schematic diagram of electrical measurement and temperature controlling apparatus.

reported that the slope is pressure sensitive and in a vacuum of 5×10^{-2} Torr the resistance doubled in less than 3000 min (i.e., approximately two days). These results suggest that coating the junction with silicon monoxide immediately after its formation tends to preserve the initial junction characteristics.[16]

The measurement techniques described below are designed to yield 0.1% precision, which is required for the analysis of the data in the following section. The stability of the junction is consequently of prime importance. After the sample has been aged for four months the junction resistance changes 0.1% in about 3000 min which is approximately the time required to take the data for the volt-current characteristics analyzed in Sec. IV.

The change in the zero-applied-voltage resistance is accompanied by a shift in the volt-current characteristics. After storing the samples for four months, the electrical characteristics up to 1.0 V are quite stable with time. The initial plot of log j versus V and the plot after aging for four months are shown in Fig. 2(a). The final plot is reproducible to within 0.1% at a constant temperature even when the sample has been cycled to liquid-nitrogen temperature between measurements. No shorting,[17] hysteresis,[17] long time constant,[17] forming,[18,19] negative dynamic resistance,[18,20] or change in the direction or rectification with temperature[15] effects were observed. A small rectification effect is observed and is discussed in Sec. IV.

Aluminum oxide films formed by a number of different techniques have been investigated in this laboratory. Only the oxides gfown thermally, as described above, were used in the experiments reported here. However, the results are considered to apply to almost all aluminum oxide films of less than 30 Å thickness grown on the surface of vacuum deposited aluminum films, provided the applied voltage is less than 0.2 V. The shape of the log j versus V plot can be correlated with the oxidation technique.[21] However, the characteristic shapes for each oxidation technique are nearly identical for applied voltage less than 0.2 V, aside from a multiplicative constant which arises from small differences in the oxide film thickness.

C. Measurement Techniques

A schematic diagram of the electrical measurement and temperature controlling apparatus is shown in Fig. 3. The apparatus is designed to maintain the sample inside a temperature-controlled light-tight chamber at any temperature between the boiling point of liquid nitrogen and room temperature. Since the substrate is a relatively poor heat conductor, it is necessary to enclose the sample in a constant temperature chamber, because the sample tends to come to radiation equilibrium with the temperature of its surroundings. The apparatus is similar to that of Ure[22] with the addition of a proportional temperature controller which regulates the dc heater power. With this apparatus the temperature can be maintained within ±0.2°K during the time required to measure the volt-current characteristics. The controlled temperature is measured with a copper-constantan thermocouple.

The dc volt-current characteristics at constant temperature are measured by setting a predetermined voltage across the junction, as compared with a volt potentiometer, and measuring the voltage drop across a 100 ohm ($\pm0.05\%$) resistance in series with the sample with a K-3 potentiometer to determine the current.

III. THEORY

The net current density j in the direction normal to the insulating film in a metal-insulator-metal system, from a conducting region 1 to a conducting region 2, through the potential barrier formed by the forbidden energy gap of the insulator, is given by Stratton's[8] general equation

$$j=\frac{4\pi mq}{c_1^2h^3}\left[\frac{\pi c_1kT}{\sin(\pi c_1kT)}\right]\exp(-b_1)[1-\exp(-c_1qV)], \quad (1)$$

where

$$b_1=\alpha\int_{x_{11}}^{x_{21}}(\varphi-\zeta_1)^{1/2}dx; \quad c_1=\frac{\alpha}{2}\int_{x_{11}}^{x_{21}}(\varphi-\zeta_1)^{-1/2}dx,$$

which are functions of the applied voltage V that arise

[16] This result is in disagreement with the results of J. Penley, J. Appl. Phys. 33, 1623 (1962), but suggests that the pressure sensitive resistance change may be due to diffusion through the unprotected top metal layer, see J. J. Wortman and R. M. Burger, Second Annual Symposium on the Physics of Failure in Electronics, 1963 (unpublished).

[17] D. Meyerhoffer and S. A. Ochs, J. Appl. Phys. 34, 2535 (1963).

[18] T. W. Hickmott, J. Appl. Phys. 33, 2669 (1962).

[19] S. R. Pollack, J. Appl. Phys. 34, 877 (1963).

[20] H. Kanter and W. A. Feibelman, J. Appl. Phys. 33, 3580 (1962).

[21] The correlation of the shape of the log j versus V curves with the oxidation technique is the subject of another paper by one of the authors (T.E.H.) (to be published).

[22] R. W. Ure, Jr., Rev. Sci. Instr. 28, 836 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

from the WKB approximation and can be computed for a given barrier $\varphi(x,V)$. The limits x_{11} and x_{21} are the classical turning points given by the values of x for which $\varphi(x,V)=\zeta_1$, ζ is the Fermi energy, x is the direction normal to the insulating film, q is the charge and m the mass of the electron, h is Planck's constant, T is the absolute temperature and $\alpha=4\pi(2m)^{1/2}/h$.

Equation (1) is derived assuming a parabolic energy-momentum relation with the same electron mass in each of the three regions. Stratton[8] has shown that the equation is still valid when the effective mass in the insulator is less than the mass in either conducting region. It is found that the more general form of Stratton's equation for a nonparabolic energy-momentum relation in the insulator is not required to explain the results in Sec. IV. This presumably arises because the barrier heights deduced in Sec. IV are a small fraction of the energy gap. Other conditions on the range of application of Eq. (1) are discussed by Stratton.

The last term in brackets in Eq. (1) represents the back current, or electrons tunneling from the conducting region 2 to conducting region 1 against the applied electric field. This term diminishes in importance rapidly with increasing applied voltage. The voltage dependence of the forward current results predominantly from $b_1(V)$ in the exponential.

In deriving the equation it is assumed that the current is largely due to electrons whose energies are close to the Fermi energy in the conducting region 1. As a result an expansion with respect to energy near ζ_1 is employed. The result of this procedure is to assume that most of the current is forward current, and that even the back current arises from electrons whose energies are close to ζ_1. This is a very good approximation for all but the very small applied voltages.

The back current is more accurately represented by an expansion about the Fermi energy in the conducting region 2 given by $\zeta_2=\zeta_1-qV$.[23] From the symmetry it follows that the current density is more accurately given by

$$j=\frac{4\pi mq}{h^3}\left\{\frac{e^{-b_1}}{c_1^2}\left[\frac{\pi c_1 kT}{\sin(\pi c_1 kT)}\right]-\frac{e^{-b_2}}{c_2^2}\left[\frac{\pi c_2 kT}{\sin(\pi c_2 kT)}\right]\right\}. \quad (2)$$

From this more accurate expression both Stratton's and Simmons'[24] expressions for the current density, assuming an arbitrary barrier shape $\varphi(x,V)$, can be derived depending on the particular assumptions employed.[25] Both Eq. (2) and Simmons'[24] Eq. (20) are intractable to experimental analysis.

The difficulty in arriving at any straightforward solution for the current density, where an arbitrary (or unknown) barrier shape $\varphi(x,V)$ is involved, is the evaluation of the integrals arising from the WKB approximation. Stratton expands the functions around energies near the Fermi level to avoid the difficulty, whereas Simmons employs an approximate value for the integrals.

The advantage of Stratton's theory is that it holds for nonzero temperatures and in the form given in Eq. (1) it is tractable to experimental analysis. The unknown functions $b_1(V)$ and $c_1(V)$ can be deduced by fitting the experimental current-density data, as functions of temperature and voltage, to Eq. (1) as shown in the following section. The disadvantage is the inaccuracy in the back current, which is not serious as will be indicated.

The disadvantages to Simmons' theory are the restriction to absolute zero temperature and the experimental intractability. The advantage is that since electrons of all energies are taken into account the back current is more accurate. By combining the two approaches, some of the advantages of both can be retained. In particular, the average barrier height as defined by Simmons and the tunnel path length can be given by

$$\bar{\varphi}-\zeta_1=b_1/(2c_1), \quad (3)$$

and

$$(\Delta x)^2=2b_1c_1/\alpha^2, \quad (4)$$

respectively, in terms of Stratton's functions which can be deduced from experiment.

Finally, it is easy to show from Eq. (2) (using the values of the functions b_1 and c_1 deduced in the following section) that at a given temperature the current does not rise as rapidly with increasing voltage as implied by Eq. (1). The error decreases monotonically with increasing voltage and at 0.10 V the correction attenuates the tunnel current only about 3%. For diminishingly small applied voltages Simmons' and Stratton's expressions (at $T=0$) are equivalent. At a constant voltage, the current is less than that predicted by Eq. (1). The error decreases monotonically from a maximum of about 2% to about 1% at room temperature. The corrections to Eq. (1) stemming from the more accurate account of the back current are seen to be small.

IV. EXPERIMENTAL RESULTS

A. Temperature Dependence

It will now be shown that the experimental current density data as a function of temperature and voltage can be fitted to the functional form of the temperature dependence in Eq. (1). The temperature dependence is given by

$$j(T)\propto[\pi c_1 kT/\sin(\pi c_1 kT)]. \quad (5)$$

The functional form of Eq. (5) arises from the Fermi distribution of electrons in the metal layers and will not fit unless the electrons tunnel directly from the metal on one side to the metal on the other side (since voltages

[23] The authors are indebted to R. Stratton for pointing out this correction to the theory.

[24] J. G. Simmons, J. Appl. Phys. 34, 1793 (1963).

[25] In deriving Simmons' (Ref. 24) Eq. (20) some differences in the β factors from those of Simmons arise; however, in practice $\beta\approx 1$ over the range of interest.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

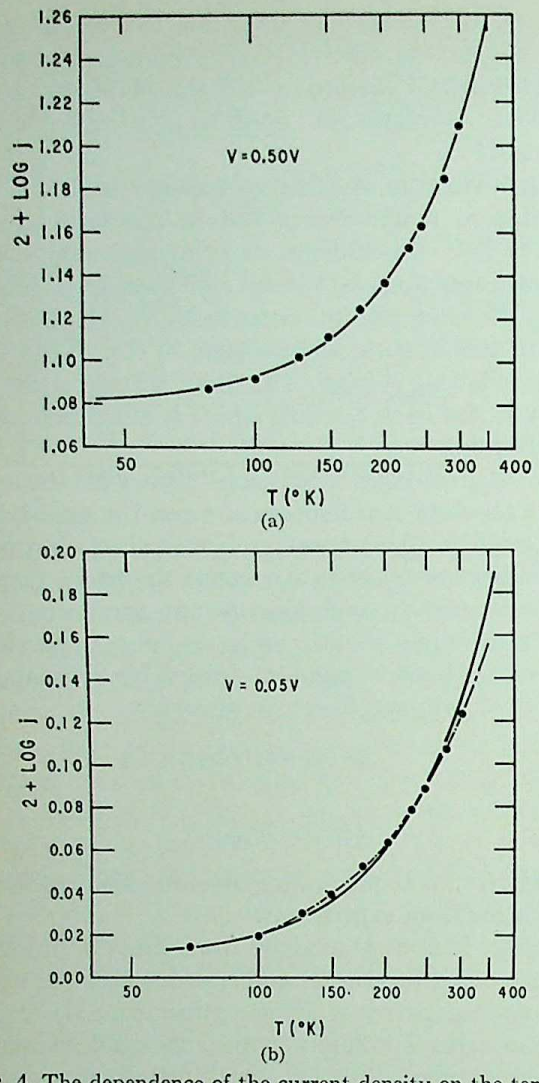


FIG. 4. The dependence of the current density on the temperature at (a) 0.50 V and (b) 0.05 V. The points are experimental data, and the curve, which is theoretical, corresponds to Eq. (1) [or the expression in (5)].

less than those required for Fowler-Nordhiem emission are employed).

The coefficients b_1 and c_1 are determined, at any particular voltage, by fitting the experimental plot of $\log_{10} j(T)$ versus $\log_{10} T$ to the theoretical plot of Eq. (5) which is $\log_{10}[\pi c_1 kT/\sin(\pi c_1 kT)]$ versus $\log_{10}(\pi c_1 kT)$. A typical example is shown in Fig. (4a). The value of c_1 is deduced from the temperature corresponding to $\pi c_1 kT=1$. Using Eq. (1), b_1 can be calculated from $j(0)$, which is deduced from the current density extraploated to absolute zero temperature corresponding to $[\pi c_1 kT/\sin(\pi c_1 kT)]=1$. The quality of the fit shown in Fig. 4(a) for 0.50 V applied potential is characteristic for all voltages between 0.10 and 0.90 V. It should be noted that the scale of the ordinate axis has been expanded greatly compared to the abscissa in order that the relatively small (e.g., compared to Schottky) temperature dependence (between 30 to 50% change in going from liquid nitrogen to room temperature) can be compared accurately with the theoretical form.

The quality of the fit of the experimental current-density data, as functions of temperature and voltage, to the functional form of the theoretical temperature dependence establishes that tunneling directly between the two aluminum layers is the predominant mechanism for the electronic conduction through the aluminum oxide film.

A typical example of the fit of the data to the expression (5) for $V<0.10$ V is shown in Fig. 4(b). The deviations may arise from small thermal emf's which can produce internal voltages, but do not appear to be due to the back current [i.e., fit the transcendental temperature dependence given in Eq. (2)]. The form of the experimental curve, shown as the dashed curve in Fig. 4(b), is characteristic of all the samples investigated at very small applied voltages.

The temperature dependence in Eq. (1) can be expanded for small temperatures in a binomial series with respect to $\pi c_1 kT$ which gives[8]

$$\Delta j = j(T) - j(0) = \frac{j(0)}{6}(\pi c_1 kT)^2 + \frac{7j(0)}{360}(\pi c_1 kT)^4 + \cdots . \quad (6)$$

Chow[26] has analyzed the temperature dependence of the current density in the Be-BeO-Au system by fitting the data to the quadratic dependence given in Eq. (6) and neglecting the higher order terms. The $j(0)$ was calculated, assuming the quadratic dependence, using two of the data points. The result of plotting $\log_{10} \Delta j = \log_{10}[j(T)-j(0)]$ versus $\log_{10} T$ is a straight line of slope 2.04 up to 300°K. Similar results for the Al-Al_2O_3-Al system have also be reported.[17] Analyzing the temperature dependence data in this way is subject

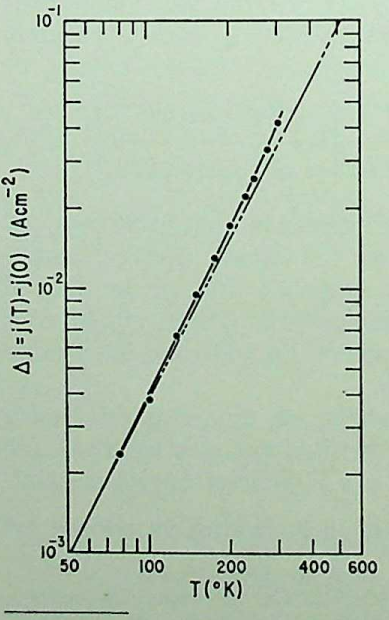


FIG. 5. The dependence of the change in the current density at 0.50 V on the temperature. The points are experimental data, the dashed line is the T^2 approximation, and the curve, which is theoretical, corresponds to Eq. (6) and includes the T^4 term.

[26] C. K. Chow, J. Appl. Phys. 34, 2919 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to considerable error. In order that the quadratic approximation hold, $j(0)$ must be determined from the data over a small range of temperatures usually between 78 and 150°K. This may result in some inaccuracy in $j(0)$ and considerable deviation of $\Delta j(T)$ from the T^2 line over the range of temperatures where the T^2 approximation is valid (cf. Chow's[26] Fig. 1). However, $j(0)$ can be determined with good accuracy using the technique described above. Using the $j(0)$ determined from Fig. 4(a), a plot of $\log_{10}\Delta j$ versus $\log_{10}T$ is shown in Fig. 5 for the experimental data shown in Fig. 4(a). Results quite similar to those of Chow are obtained except that good agreement with the T^2 linear approximation (shown as the dashed line in Fig. 5) is obtained for temperatures less than 100°K. The solid curve is a plot of the theoretical temperature dependence as given by Eq. (6) including the T^4 term, using the value of $c_1=16.0$ eV^{-1} determined from Fig. 4(a). A straight line of slope slightly greater than two can be fitted to the data in Fig. 5, but in fact it is seen that the deviation of the experimental data from the T^2 line is due to the T^4 term in Eq. (6).

B. Functions $b_1(V)$ and $c_1(V)$

The coefficients b_1 and c_1 are functions of voltage and arise from the WKB approximation. The specific volt-current characteristic for a given barrier potential $\varphi(x,V)$ can be given from Eq. (1) by evaluating b_1 and c_1 as functions of V. Stratton[8] has developed Eq. (1) further for small applied voltages, while maintaining an arbitrary barrier shape, by expanding b_1 and c_1 as the following quadratic power series in qV

$$b_1=b_{10}-b_{11}(qV)+b_{12}(qV)^2, \tag{7}$$

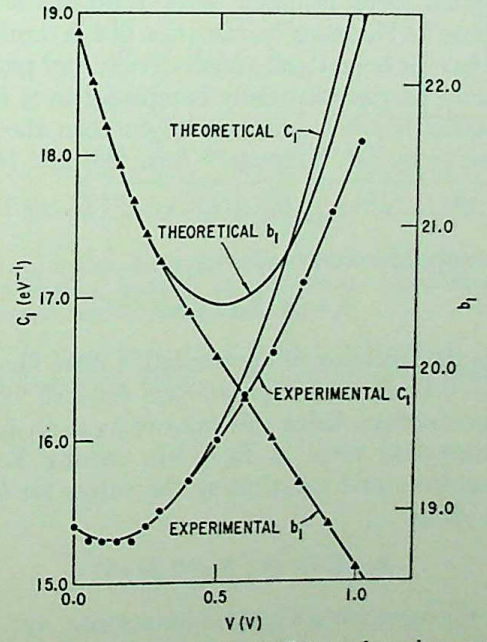


FIG. 6. The dependence of the coefficients b_1 and c_1 on voltage. The points are deduced from the experimental data by fitting the temperature dependence. The theoretical curves are quadratic power-series expansions corresponding to Eqs. (7) and (8).

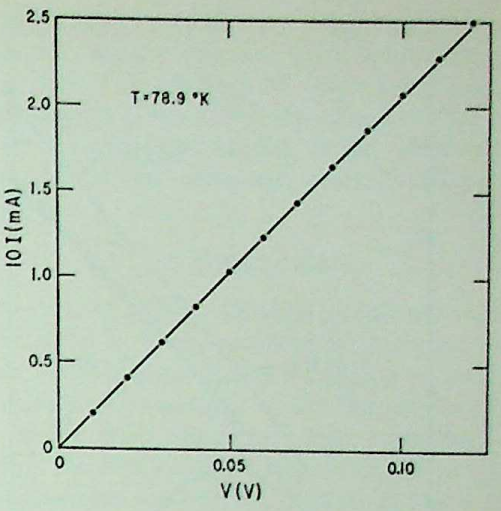


FIG. 7. Low-voltage Ohmic region. The current as a function of small applied voltage at $T=78.9$°K.

and

$$c_1=c_{10}-c_{11}(qV)+c_{12}(qV)^2. \tag{8}$$

By fitting the experimental current density data to the theoretical temperature dependence for a set of applied voltages in the range $0\leq V\leq 1.0$ V, b_1 and c_1 can be deduced as functions of V. The results are shown in Fig. 6. The experimental curves are plots of b_1 and c_1 deduced from the experimental data. The theoretical curves are the quadratic power-series expansions given in Eqs. (7) and (8) which have been fitted to the experimental curves for small applied voltages. The coefficients for the power series are listed in Table II. The region of applied voltage for which the quadratic power-series expansions in V, of the integrals arising from the WKB approximation, are valid, is seen to be $0.0\leq V\leq 0.3$ V.

Although c_1 can be represented approximately by the low-voltage quadratic over a substantially larger range of applied voltage, b_1 cannot. A plot of the current density as a function of voltage [cf. Fig. 2(a)] suggests a nearly linear approximation for b_1 in the range $0.2<V<1.0$ V, preceded by an Ohmic region (cf. Fig. 7). Stratton,[8] in analyzing the earlier data of Fischer and Giaever[27] and Pimbley,[28] fitted the quadratic power series for b_1 out to about 1.4 V. This procedure is equivalent essentially to fitting a parabola to the lower part of the experimental curve for b_1 shown in Fig. 6, which does not converge to the experimental curve for b_1 for

TABLE II. Quadratic power-series coefficients.

b_1	c_1 (eV)$^{-1}$
$b_{10}=22.37$	$c_{10}=15.38$ (eV)$^{-1}$
$b_{11}=7.56$ (eV)$^{-1}$	$c_{11}=1.32$ (eV)$^{-2}$
$b_{12}=7.4$ (eV)$^{-2}$	$c_{12}=5.29$ (eV)$^{-3}$

[27] J. C. Fisher and I. Giaever, J. Appl. Phys. 32, 172 (1961).
[28] W. T. Pimbley (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

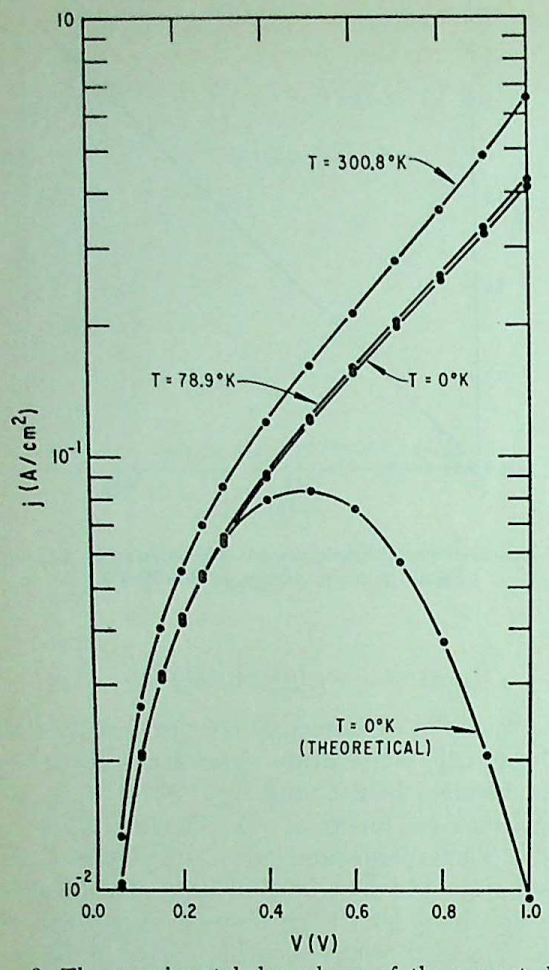


FIG. 8. The experimental dependence of the current density on the voltage at room, liquid nitrogen and extrapolated to absolute zero temperature. The theoretical curve was calculated, assuming the quadratic power-series approximations for b_1 and c_1, using Eq. (1).

small V. The coefficients of the quadratic expansions are related to physical properties of the potential barrier.[8] The numerical values for the physical properties reported by Stratton, therefore, are in error.

The low-voltage quadratic expansions given in Eqs. (7) and (8) using the coefficients in Table II were used to calculate the theoretical volt-current characteristic for $T=0$°K shown in Fig. 8. The experimental characteristics for room, liquid nitrogen and extrapolated to absolute zero temperature are also shown. It is apparent that the theoretical curve is a good representation for $0\leq V\leq 0.3$ V, which is the range of voltage for which the quadratic power-series expansion for b_1 is valid, and that the divergence between the theoretical and experimental curve for $T=0$°K results from the functional form of b_1 shown in Fig. 6.

The functional form of b_1 and c_1 shown in Fig. 6 for $V>0.3$ V is not expected to be universal. In particular, the functional form shown here pertains to an oxide grown by the technique described in Sec. IIA.

C. Physical Properties of the Barrier Layer

The numerical values of the coefficients in Table II can be related to the physical properties of the barrier when a particular barrier shape is assumed. Using Simmons' expression for the tunnel path length given in Eq. (4), b_1 and c_1 given by Eqs. (7) and (8), and the coefficients from Table II, it is easy to show that Δx is approximately constant over the range $0.2<V<0.6$ V which is consistent with a trapezoidal (or rectangular) barrier shape.

Using the asymmetry in the experimental current density with respect to the polarity of the applied potential, and the relation for asymmetrical barriers that[8]

$$b_{11}+b_{21}=c_{10}$$

a value of $b_{21}=7.82$ (eV)$^{-1}$ is obtained.

Assuming a trapezoidal barrier shape for $\varphi(x,V)$, integrating, carrying out the Taylor expansions with respect to energy and the binomial expansion in Eq. (7) (and assuming a corresponding series for b_2) leads to the following relationships between the physical parameters and the coefficients:

$$\left.\begin{aligned} b_{10}&=\frac{2\alpha L}{3}\left(\frac{\chi_2'^{3/2}-\chi_1'^{3/2}}{\chi_2'-\chi_1'}\right),\\ b_{11}&=(\alpha L\chi_2'^{1/2}-b_{10})/(\chi_2'-\chi_1'),\\ b_{21}&=-(\alpha L\chi_1'^{1/2}-b_{10})/(\chi_2'-\chi_1'),\end{aligned}\right\} \tag{9}$$

where

$$\chi_n=\chi_n'+\Delta\chi;\ n=1,2. \tag{10}$$

χ_1' and χ_2' are the effective heights of the trapezoidal barrier above the Fermi level on the side adjacent to metal 1 and 2, respectively, which differ from the true metal-insulator work functions χ_1 and χ_2 by an amount $\Delta\chi$. It will be assumed that the barrier is a result of two unequal metal-insulator work functions and the image force. In Simmons'[29] evaluation of b in terms of an average barrier height and an effective tunnel path, the image force barrier essentially is replaced by a rectangular barrier which is lower in height than the image force barrier by [cf. Simmons'[29] Eqs. (33) and (40)]

$$\Delta\chi=[1.15\lambda L/(x_2-x_1)]\ln[x_2(L-x_1)/x_1(L-x_2)] \tag{11}$$

for zero applied voltage, where

$$\lambda=(q^2/8\pi\kappa L)\ln 2,$$

and L is the thickness of the insulating film. The quantity $\Delta\chi=0.11$ eV for the parameters deduced from the experimental data. Using this value of $\Delta\chi$ as an approximate correction term in Eq. (10), solving Eqs. (9) simultaneously and substituting the values for b_{10}, b_{11}, and b_{21} yields

$$\chi_1=0.78\text{ eV};\ \chi_2=0.89\text{ eV}$$

which will represent a slight underestimate, and

$$L=25.7\text{ Å}$$

[29] J. G. Simmons, J. Appl. Phys. 34, 2581 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

assuming an effective electron mass in the insulator $m^*/m=1$. Meyerhoffer and Ochs[17] have pointed out that the effective mass theoretically should be near to one.[30,31] Assuming a dielectric constant of 8.0,[18] the thickness of the insulating layer calculated from capacitance is 25.7 Å, and the tunnel path length, calculated from Eq. (4), is 25.6 Å. The aluminum film corresponding to χ_1 is the one on which the oxide was grown. Greater currents are passed when this layer is positive, which is in agreement with other observations.[15,17]

A plot of the experimental data for a thicker aluminum oxide (formed in an oxygen glow discharge) and larger applied voltages is shown in Fig. 9. The data are plotted as log I/V^2 versus $1/V$ for comparison with the Fowler-Nordheim theory.[2] Including the image force correction the slope (for large V) results in a value of $\chi_2=0.85$ eV, which is in good agreement with the value deduced from the low-voltage measurements.

[*Note added in proof.* Recent investigations on the thicker aluminum oxides at larger applied voltages (similar to specimens used to obtain the data shown in Fig. 9) indicate considerably larger effective barrier heights and smaller effective thicknesses than are obtained from the analysis of low-voltage data. The barrier heights given above are based on the assumption of a trapezoidal barrier shape which does not appear to be an adequate representation for all voltages.]

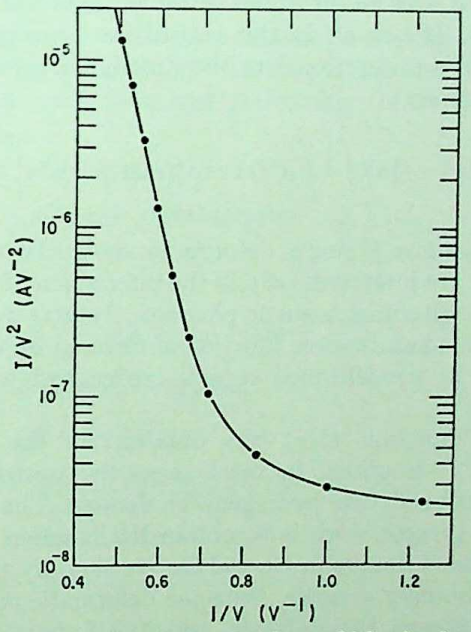


FIG. 9. A Fowler-Nordheim plot for a thick aluminum oxide film. The slope leads to an approximate metal-insulator work function of $\chi_2=0.85$ eV.

[30] E. O. Kane, Phys. Chem. Solids 1, 245 (1957).
[31] R. Braunstein and E. O. Kane, Phys. Chem. Solids 23, 1423 (1962).

V. CONCLUSIONS

The current density through aluminum oxide films less than 30 Å thick, sandwiched between two aluminum films, is presented as a function of temperature and voltage. The analysis of the temperature dependence of the data established that tunneling directly between the two aluminum layers is the predominant mechanism for electron conduction through the aluminum oxide film.

The data are analyzed using the theory of Stratton. The theory is modified to take more accurate account of the back current and compared with the theory of Simmons. Both the modified and Simmons expressions are intractable to experimental analysis. It is indicated that the errors involved in using the simplified version of Stratton's theory are small, therefore, the original form given in Eq. (1), is employed.

The region of validity is established, for quadratic power-series expansions in V, of the integrals arising from the WKB approximation. Fitting the quadratics to the experimental curves for these integrals results in approximate values for the heights of the potential barrier above the Fermi level and the barrier thickness, when a trapezoidal barrier shape (modified by image forces) is assumed.

The actual shape of the potential barrier cannot be extracted with any degree of confidence using only the low-voltage current-density data as a function of temperature. Therefore, even though the preliminary Fowler-Nordheim analysis results in a barrier height in agreement with the value obtained for an image force barrier, a more elaborate investigation is required.

ACKNOWLEDGMENTS

The authors acknowledge with thanks the many helpful and stimulating discussions with R. Stratton and N. G. Einspruch, the technical assistance of P. J. May, Jr., J. W. Taylor for the preparation of samples, and J. C. Blair for the data shown in Fig. 9.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Interaction of Electron on the Donor Center with the Crystal Lattice in Germanium and Silicon

ANDRZEJ MYSZKOWSKI AND STANISŁAW GÓMUŁKA
Institute of Theoretical Physics, Warsaw University, Warsaw, Poland

(Received 8 July 1963)

The diagonal matrix element of interaction of the bound electron with longitudinal acoustic phonon, squared and averaged over all directions of the phonon wave vector, was calculated for n-type Ge and Si assuming a single-deformation-potential constant. The donor wave functions have been taken in the elipsoidal form. Some simple integrals containing the squared and averaged matrix element are given and are used, as an example, to get the one-phonon transition rate in the impurity conduction by "hopping" and the relaxation energy of the lattice deformed by interaction with the donor electron. Appreciable differences are found between the case of elipsoidal and spherical donor wave functions.

INTRODUCTION

THE electron bound to the donor center in the homopolar crystal produces a deformation (a change of the equilibrium positions of the crystal atoms) in the vicinity of the center because of the interaction with acoustic phonons. This deformation influences several physical processes in the crystal. It changes the one-phonon transition probability and gives the possibility of the multiphonon processes in the case of the impurity conduction by "hopping."[1–4] A similar effect occurs in the case of the thermal recombination of donors and electrons.[5] The deformation of the crystal lattice gives also the difference between the energies of thermoionization and photoionization of donors.[6] In the process of absorption of the electromagnetic wave by the bound electron this deformation gives the possibility of the simultaneous absorption of photon and emission of phonons and so has an influence on the shape of the absorption lines.[7] In all papers referred to above[1–7] taking into account lattice-deformation effects requires computation of the diagonal matrix element of deformation potential of the phonon in the ground state of the electron on the isolated donor. This matrix element squared must then be averaged over all directions of the wave vector τ of the phonon. The result in the case of the isotropic sound velocity is proportional to the function $U(\tau)$ which is by definition

$$U(\tau)=(\tfrac{1}{4}\pi)\int\left|\int u^*(\mathbf{r})\exp(i\tau\boldsymbol{\Omega}\mathbf{r})u(\mathbf{r})d^3x\right|^2 d\Omega, \quad (1)$$

if we assume a single deformation-potential constant and we are interested only in the interaction of electron with longitudinal acoustic phonons. By $u(\mathbf{r})$ we denote the ground-state wave function of electron on impurity center in nondeformed crystal lattice and we define $\tau=\tau\boldsymbol{\Omega}$.

The function $U(\tau)$ was obtained in the referred papers[1–7] assuming in most cases the isotropic conduction band (the hydrogen-like donors). The purpose of the present work is to obtain the function $U(\tau)$ in the case of donors in Ge and Si. We simplify the problem assuming a single, isotropic deformation-potential constant and the isotropic velocity of the sound and taking into account only longitudinal acoustic phonons. The results of this paper can be easily used in another materials with similar shape of the conduction-band minima. The paper contains also same simple integrals with the function $U(\tau)$ for Ge and Si. We used them, as an example, to get the corrections to the results of the Ref. 2 which deal with the one-phonon transition rate in the impurityc onduction by "hopping." We obtain also corrections to the lattice relaxation energy.

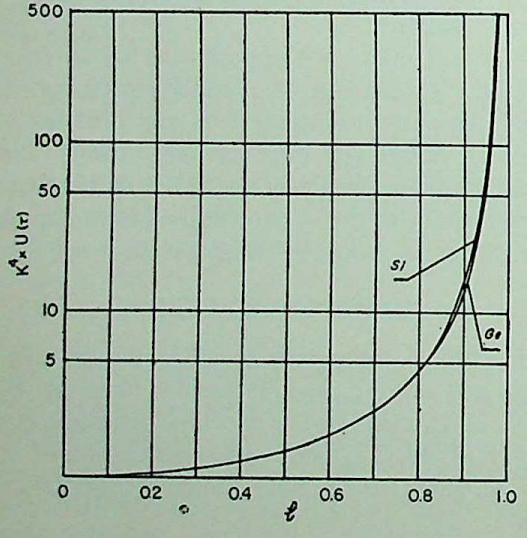


FIG. 1. The dependence of the function $U(\tau)$ on $l=[\alpha/(\alpha+1)]^{1/2}\times[1+(4/a^2\tau^2)]^{-1/2}$. For $l\to 1$ the asymptotic behavior is $(1-l^2)^{-3}/24$ and $(1-l^2)^{-3}/18$ for Ge and Si, respectively.

CALCULATION OF THE FUNCTION $U(\tau)$

In Ge and Si we have in the presence of the chemical shift the donor ground-state wave functions of the

[1] J. Mycielski, Phys. Rev. 125, 46 (1962).
[2] J. Mycielski, Phys. Rev. 125, 1975 (1962).
[3] J. Yamashita and T. Kurosawa, J. Phys. Soc. Japan 15, 802 (1960).
[4] N. F. Mott, and W. D. Twose, Advan. Phys. 10, 107 (1961).
[5] H. Gummel and M. Lax, Ann. Phys. (N. Y.) 2, 28 (1957).
[6] M. F. Deĭgen, Zh. Eksperim. i Teor. Fiz. 31, 504 (1956) [English transl.: Soviet Phys.—JETP 4, 424 (1957)].
[7] M. Lax and E. Burstein, Phys. Rev. 100, 592 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

form[8]

$$u(\mathbf{r})=n^{-1/2}\sum_{q=1}^{n}F_q(\mathbf{r})\phi_q(\mathbf{r}). \tag{2}$$

The sum is performed over the n minima of the conduction band (n equals 4 and 6 in Ge and Si, respectively). ϕ_q is the Bloch function for the qth minimum and F_q is of the form

$$F_q(x,y,z)=(\pi a^2 b)^{-1/2}\times\exp\{-[(x^2+y^2)/a^2+z^2/b^2]^{1/2}\}, \tag{3}$$

where the z axis is parallel to the axis containing the qth minimum. The constants a and b are the transverse and longitudinal radii of the orbit, respectively.

With (2) we have

$$\int u^*(\mathbf{r})\exp(i\tau\Omega\mathbf{r})u(\mathbf{r})d^3x=n^{-1}\int\sum_{p,q=1}^{n}F_pF_q\phi_p^*\phi_q\exp(i\tau\Omega\mathbf{r})d^3x. \tag{4}$$

Neglecting terms with $p\neq q$ because of the rapid oscillation of $\exp[i\mathbf{r}(\boldsymbol{\kappa}_q-\boldsymbol{\kappa}_p)]$, where $\boldsymbol{\kappa}_q$ is the wave vector for the qth minimum, and performing the integration we obtain

$$\int u(\mathbf{r})\exp(i\tau\Omega\mathbf{r})u(\mathbf{r})d^3x=n^{-1}\sum_{q=1}^{n}f_q, \tag{5}$$

where

$$f_q=K^{-2}[1-l^2(\Omega\mathbf{p}_q)^2]^{-2}. \tag{6}$$

We denoted here $\boldsymbol{\kappa}_q=\kappa_q\mathbf{p}_q$ and

$$K=1+\tfrac{1}{4}a^2\tau^2, \tag{7}$$

$$l^2=[\alpha/(\alpha+1)][1+(4/a^2\tau^2)]^{-1}, \tag{8}$$

$$\alpha=(a/b)^2-1. \tag{9}$$

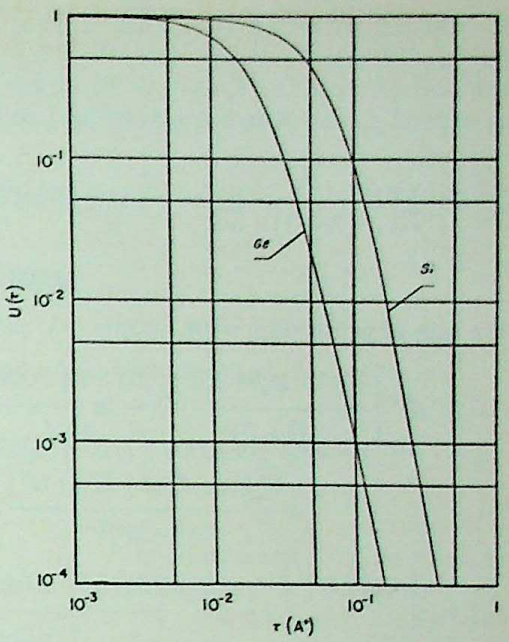


FIG. 2. The function $U(\tau)$ versus wave number τ of the phonon. For $\tau\to\infty$ the asymptotic behavior is $367.4\times K^{-4}$ and $14.81\times K^{-4}$ for Ge and Si, respectively.

For arbitrary τ we have $0\leq l^2<1$.

From (1) and (5) we obtain

$$U(\tau)=n^{-2}\sum_{p,q=1}^{n}u_{pq}(\tau), \tag{10}$$

where

$$u_{pq}(\tau)=(4\pi)^{-1}\int f_pf_qd\Omega. \tag{11}$$

Performing the integral in (11) and denoting $\mathbf{p}_p\mathbf{p}_q=(p,q)$ we obtain

$$u_{pq}=(4\pi K^4)^{-1}\{M[l,(p,q)]+M[l,-(p,q)]\}, \tag{12}$$

where

$$M[l,(p,q)]=(\pi/2)\frac{[(2-l^2)^2-l^4(p,q)^2]\{(7-l^2)[1-(p,q)^2]-2\}+6l^4(p,q)^2[1-(p,q)^2]}{[1-(p,q)^2](1-l^2)[(2-l^2)^2-l^4(p,q)^2]^2}$$
$$+(\pi/16)\frac{\{[18-l^2-l^2(p,q)][1-(p,q)]+4\}[2-l^2-l^2(p,q)]+3l^4[1-(p,q)][1+(p,q)]^2}{[1-(p,q)]^{3/2}[2-l^2-l^2(p,q)]^{5/2}}$$
$$\times\ln\frac{(1+l)\{[1-l(p,q)]\{[2-l^2-l^2(p,q)]^{1/2}+l[1-(p,q)]^{1/2}\}+\{(1-l^2)[1-(p,q)]^{1/2}\}\}}{(1-l)\{[1+l(p,q)]\{[2-l^2-l^2(p,q)]^{1/2}-l[1-(p,q)]^{1/2}\}+\{(1-l^2)[1-(p,q)]^{1/2}\}\}}. \tag{13}$$

In the case of $(p,q)=\pm1$, u_{pq} reduces to

$$\frac{33-40l^2+15l^4}{48(1-l^2)^3}+\frac{5}{32l}\ln\frac{1+l}{1-l}=F(4,\tfrac{1}{2};\tfrac{3}{2};l^2). \tag{14}$$

[8] W. Kohn, in *Solid-State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5, p. 257.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

For $\tau=0(l=0)$, we have from (12) and (13) $u_{pq}=1$ for all p and q and so $U(\tau)=1$. For $\tau\to\infty$, $(l\to[\alpha/(\alpha+1)]^{1/2})$ we have $u_{pq}\to 0$ and $U(\tau)\to 0$.

In the isotropic case ($a=b$, $\alpha=0$, $l=0$) we have simply $u_{pq}=K^{-4}$ and $U(\tau)=K^{-4}$.

If we expand f_p and f_q in powers of l^2 and perform the integral in (11) we get instead of (12)

$$u_{pq}=K^{-4}\sum_{r=0}^{\infty}l^{2r}\frac{1}{(2r+1)!!}\sum_{s=0}^{r}(r-s+1)(s+1)\times\sum_{t=0}^{s}\binom{2s}{2t}(p,q)^{2t}[1-(p,q)^2]^{s-t}(2r-2s+2t-1)!!(2s-2t-1)!!. \quad (15)$$

GERMANIUM AND SILICON

In the case of germanium $n=4$, $(p,q)=+1,\pm\frac{1}{3}$. From (10), (12) and (13) we obtain

$$U_{\mathrm{Ge}}(\tau)=2^{-8}K^{-4}\left\{\frac{4(33-40l^2+15l^4)}{3(1-l^2)^3}+\frac{10}{l}\ln\frac{1+l}{1-l}+\frac{27[(3-2l^2)(3-l^2)(19-4l^2)+6l^4]}{(1-l^2)(3-2l^2)^2(3-l^2)^2}\right.$$
$$+\frac{9[(18-l^2)(3-2l^2)+6l^4]}{l(3-2l^2)^{5/2}}\ln\frac{(1+l)\{(3-l)[(3-2l^2)^{1/2}+l]+3(1-l^2)\}}{(1-l)\{(3+l)[(3-2l^2)^{1/2}-l]+3(1-l^2)\}}$$
$$\left.+\frac{9[(63-2l^2)(3-l^2)+6l^4]}{2^{5/2}l(3-l^2)^{5/2}}\ln\frac{(1+l)\{(3+l)[2^{1/2}(3-l^2)^{1/2}+2l]+6(1-l^2)\}}{(1-l)\{(3-l)[2^{1/2}(3-l^2)^{1/2}-2l]+6(1-l^2)\}}\right\} \quad (16)$$

or from (10) and (15)

$$U_{\mathrm{Ge}}(\tau)=K^{-4}(1.000+1.333l^2+1.644l^4+1.943l^6+2.247l^8+2.567l^{10}+3.191l^{12}+3.285l^{14}+3.608l^{16}+\cdots). \quad (17)$$

In the case of silicon $n=6$, $(p,q)=\pm1$, 0. From (10), (12) and (13) we obtain

$$U_{\mathrm{Si}}(\tau)=2^{-4}3^{-2}K^{-4}\left\{\frac{33-40l^2+15l^4}{(1-l^2)^3}+\frac{15}{2l}\ln\frac{1+l}{1-l}+\frac{24(5-l^2)}{(1-l^2)(2-l^2)^2}\right.$$
$$\left.+\frac{3[(22-l^2)(2-l^2)+3l^4]}{l(2-l^2)^{5/2}}\ln\frac{(1+l)[(2-l^2)^{1/2}+l+1-l^2]}{(1-l)[(2-l^2)^{1/2}-l+1-l^2]}\right\}, \quad (18)$$

or from (10) and (15)

$$U_{\mathrm{Si}}(\tau)=K^{-4}(1.000+1.333l^2+1.644l^4+1.943l^6+2.263l^8+2.628l^{10}+3.048l^{12}+3.528l^{14}+4.069l^{16}+\cdots). \quad (19)$$

Dependence of $K^4U(\tau)$ on l for Ge- and Si-type materials is plotted on the Fig. 1. Taking $\alpha=18.8$, $a=70.8$ Å for Ge and $\alpha=4.2$, $a=22.1$ Å for Si we obtain the dependences of $U(\tau)$ on τ given on the Fig. 2.

PARTICULAR INTEGRALS

We calculate here the integrals of the type $\int_0^\infty U(\tau)\tau^\kappa d\tau$ for $\kappa=0$, 1, 2, 3. We use formula (10) and u_{pq} in the form (14) for $(p,q)=\pm1$ or in the form (15) for $(p,q)\neq\pm1$.

In the case of germanium we obtain

$$\int_0^\infty U_{\mathrm{Ge}}(\tau)d\tau=\frac{15\pi}{64a}\left\{\frac{1}{3}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\arcsin\left[\frac{\alpha}{\alpha+1}\right]^{1/2}+1.000+0.1667\frac{\alpha}{\alpha+1}+0.05722\left[\frac{\alpha}{\alpha+1}\right]^2\right.$$
$$+0.02560[\alpha/(\alpha+1)]^3+0.01327[\alpha/(\alpha+1)]^4+0.007581[\alpha/(\alpha+1)]^5$$
$$\left.+0.005640[\alpha/(\alpha+1)]^6+0.002991[\alpha/(\alpha+1)]^7+\cdots\right\}, \quad (20)$$

$$\int_0^\infty U_{\mathrm{Ge}}(\tau)\tau d\tau=\frac{1}{2a^2}\left\{\frac{1}{6}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\ln\frac{1+[\alpha/(\alpha+1)]^{1/2}}{1-[\alpha/(\alpha+1)]^{1/2}}+1.000+0.3333\frac{\alpha}{\alpha+1}+0.1526\left[\frac{\alpha}{\alpha+1}\right]^2\right.$$
$$+0.08190[\alpha/(\alpha+1)]^3+0.04855[\alpha/(\alpha+1)]^4+0.03081[\alpha/(\alpha+1)]^5$$
$$\left.+0.02500[\alpha/(\alpha+1)]^6+0.01428[\alpha/(\alpha+1)]^7+\cdots\right\}, \quad (21)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$\int_0^\infty U_{\mathrm{Ge}}(\tau)\tau^2 d\tau=\frac{3\pi}{16a^3}\left\{\frac{1}{3}(\alpha+1)^{1/2}+1.000+0.5000\frac{\alpha}{\alpha+1}+0.2861\left[\frac{\alpha}{\alpha+1}\right]^2+0.1792\left[\frac{\alpha}{\alpha+1}\right]^3\right.$$
$$+0.1195[\alpha/(\alpha+1)]^4+0.08340[\alpha/(\alpha+1)]^5+0.07332[\alpha/(\alpha+1)]^6$$
$$\left.+0.04487[\alpha/(\alpha+1)]^7+0.03187[\alpha/(\alpha+1)]^8+\cdots\right\}, \quad (22)$$

$$\int_0^\infty U_{\mathrm{Ge}}(\tau)\tau^3 d\tau=a^{-4}\left\{\frac{1}{6}(\alpha+1)+\frac{1}{12}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\ln\frac{1+[\alpha/(\alpha+1)]^{1/2}}{1-[\alpha/(\alpha+1)]^{1/2}}+1.000+0.6667\frac{\alpha}{\alpha+1}+0.4578\left[\frac{\alpha}{\alpha+1}\right]^2\right.$$
$$+0.3276[\alpha/(\alpha+1)]^3+0.2427[\alpha/(\alpha+1)]^4+0.1848[\alpha/(\alpha+1)]^5+0.1750[\alpha/(\alpha+1)]^6$$
$$\left.+0.1142[\alpha/(\alpha+1)]^7+0.08591[\alpha/(\alpha+1)]^8+\cdots\right\}. \quad (23)$$

In the case of silicon we obtain

$$\int_0^\infty U_{\mathrm{Si}}(\tau)d\tau=\frac{5\pi}{24a}\left\{\frac{1}{2}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\arcsin\left[\frac{\alpha}{\alpha+1}\right]^{1/2}+1.000+0.1667\frac{\alpha}{\alpha+1}+0.05500\left[\frac{\alpha}{\alpha+1}\right]^2+0.02321\left[\frac{\alpha}{\alpha+1}\right]^3\right.$$
$$\left.+0.01133[\alpha/(\alpha+1)]^4+0.006138[\alpha/(\alpha+1)]^5+0.003602[\alpha/(\alpha+1)]^6+\cdots\right\}, \quad (24)$$

$$\int_0^\infty U_{\mathrm{Si}}(\tau)\tau d\tau=\frac{4}{9a^2}\left\{\frac{1}{4}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\ln\frac{1+[\alpha/(\alpha+1)]^{1/2}}{1-[\alpha/(\alpha+1)]^{1/2}}+1.000+0.3333\frac{\alpha}{\alpha+1}+0.1467\left[\frac{\alpha}{\alpha+1}\right]^2\right.$$
$$\left.+0.07429[\alpha/(\alpha+1)]^3+0.04145[\alpha/(\alpha+1)]^4+0.02494[\alpha/(\alpha+1)]^5+0.01597[\alpha/(\alpha+1)]^6+\cdots\right\}, \quad (25)$$

$$\int_0^\infty U_{\mathrm{Si}}(\tau)\tau^2 d\tau=\frac{\pi}{6a^3}\left\{\tfrac{1}{2}(\alpha+1)^{1/2}+1.000+0.5000\frac{\alpha}{\alpha+1}+0.2750\left[\frac{\alpha}{\alpha+1}\right]^2+0.1625\left[\frac{\alpha}{\alpha+1}\right]^3\right.$$
$$\left.+0.1019[\alpha/(\alpha+1)]^4+0.06752[\alpha/(\alpha+1)]^5+0.04683[\alpha/(\alpha+1)]^6+\cdots\right\}, \quad (26)$$

$$\int_0^\infty U_{\mathrm{Si}}(\tau)\tau^3 d\tau=\frac{8}{9a^4}\left\{\tfrac{1}{4}(\alpha+1)+\frac{1}{8}\left[\frac{\alpha}{\alpha+1}\right]^{-1/2}\ln\frac{1+[\alpha/(\alpha+1)]^{1/2}}{1-[\alpha/(\alpha+1)]^{1/2}}+1.000+0.6667\frac{\alpha}{\alpha+1}+0.4400\left[\frac{\alpha}{\alpha+1}\right]^2\right.$$
$$\left.+0.2971[\alpha/(\alpha+1)]^3+0.2073[\alpha/(\alpha+1)]^4+0.1497[\alpha/(\alpha+1)]^5+0.1118[\alpha/(\alpha+1)]^6+\cdots\right\}. \quad (27)$$

Taking $\alpha=18.8$, $a=70.8$ Å for Ge and $\alpha=4.2$, $a=22.1$ Å for Si we obtain values given in the Table I.

CORRECTIONS TO THE ONE-PHONON "HOPPING" PROBABILITY AND LATTICE RELAXATION ENERGY

According to Ref. 2 the total one-phonon transition rate for the electron transfer from the donor a to the donor b is

$$W_{ba}{}^{1\mathrm{tot}}=W_{ba}{}^{1}(1-A)^2\exp(-\alpha), \quad (28)$$

where $W_{ba}{}^{1}$ is the one-phonon transition rate if we neglect the dependence of the equilibrium positions of the lattice atoms on the state of the carrier. α and A are the quantities described in Ref. 2. In particular for low absolute temperature T

$$\alpha_{T\to 0}=\frac{E_1^2}{\pi\hbar v^3 d}\int_0^\infty U(\tau)\tau d\tau, \quad (29)$$

$$A_{T\to 0}=\frac{\hbar E_1^2}{4\pi^3 v\Delta E_{ba}{}^2 d}\int_0^\infty U(\tau)\tau^2 d\tau. \quad (30)$$

E_1 is the deformation-potential constant, ΔE_{ba} the

TABLE I. Values of the integrals of the form $\int_0^\infty U(\tau)\tau^k d\tau$ for n-type Ge and Si.

Material	$k=0$ (cm^{-1})	$k=1$ (cm^{-2})	$k=2$ (cm^{-3})	$k=3$ (cm^{-4})
n–Ge	1.78×10^{6}	2.4×10^{12}	6.1×10^{18}	2.5×10^{25}
n–Si	5.36×10^{6}	2.05×10^{13}	1.4×10^{20}	1.4×10^{27}

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Corrections to the one-phonon "hopping" probability and lattice relaxation energy for n-type Ge and Si.

Material	$\alpha_{T\to 0}$	$A_{T\to 0}$	$(W_{ba}{}^{1\mathrm{tot}}/W_{ba}{}^{1})_{T\to 0}$	E_{LR} (eV)
n-Ge	0.059(0.08)	0.067(0.08)	0.82(0.78)	$2.5\times10^{-5}(3.3\times10^{-5})$
n-Si	0.43(0.56)	0.39(0.59)	0.24(0.10)	$0.80\times10^{-3}(1.1\times10^{-3})$

difference of the electrostatic energy of the carrier in the two donors, v the velocity of sound in the crystal, and d the density of the crystal.

In the Table II we give the values of $\alpha_{T\to 0}$, $A_{T\to 0}$ and $(W_{ba}{}^{1\mathrm{tot}}/W_{ba}{}^{1})_{T\to 0}$ obtained from (28)–(30) and from data given in Table I for n-type Ge and Si samples studied in Ref. 2. We give also the values of relaxation energy E_{LR} of the lattice deformed by interaction with the donor electron, which is[2]

$$E_{\mathrm{LR}}=\frac{E_1{}^2}{4\pi v^2 d}\int_0^\infty U(\tau)\tau^2 d\tau. \tag{31}$$

In Table II we give in brackets for comparison the values obtained previously[2] in approximation of hydrogen-like, spherical donor.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Professor M. Suffczyński for his interest in this work and to Dr. J. Mycielski for suggesting the problem and for valuable suggestions and remarks.

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Measurement of the Conductivity Effective Mass in Semiconductors Using Infrared Reflection

HENRY A. LYDEN*

Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 21 November 1963; revised manuscript received 7 February 1964)

There exist various approximate relationships between the conductivity effective mass of free carriers m_c and the angular frequency ω_0 of the reflectivity minimum in the infrared arising from the free-carrier dispersion. A detailed analysis of the reflectivity equation shows that it is possible to obtain a relationship between m_c and ω_0 when in addition one has a knowledge of the free-carrier concentration, the drift mobility, and the dielectric constant of the semiconductor at very high frequencies. This relationship is shown to yield a value of $m_c{}^*=0.145$ for n-type germanium at room temperature using data presented in the literature. This value compares well with the value of $m_c{}^*=0.15$ obtained using a combination of reflection and transmission measurements. The experimental work required to obtain accurate values for m_c using this relationship is considerably less than that required by other techniques, such as cyclotron resonance, magnetoplasma, and combined infrared reflection and transmission measurements, which are presently used. Moreover, this procedure can be applied over very broad ranges of temperature and of free-carrier concentration.

INTRODUCTION

IN this paper[1] we are concerned with the determination of the conductivity effective mass, which is understood to be the effective mass entering in the free-carrier mobility

$$\mu=(e/m_c)\langle\tau\rangle, \tag{1}$$

where e is the electronic charge, m_c is the conductivity effective mass, and $\langle\tau\rangle$ is the relaxation time averaged over all of the free current carriers. The determination of this parameter has received increasing attention in recent years. The earliest and the most extensive investigations to date have utilized cyclotron resonance techniques. For example, the conductivity effective mass has been measured at liquid-helium temperature, where the relaxation time of free carriers is sufficiently long for both germanium and silicon.[2,3] Recently infrared techniques also have been employed for this determination. In one instance, the conductivity effective masses in germanium, silicon, and indium antimonide were obtained from a combination of reflection and transmission measurements conducted at room temperature.[4] In another instance only reflection measurements were required, in conjunction with varying intensities of magnetic field, to determine the conductivity effective masses of mercury selenide and indium antimonide.[5]

We are proposing in this paper that it is possible to determine m_c very accurately simply from a measure-

---

* Presently with Energy Conversion, Inc., Cambridge, Massachusetts.

[1] This paper is based upon a dissertation by the author submitted in partial fulfillment of the degree of Doctor of Science to the Electrical Engineering Department of the Massachusetts Institute of Technology and published as Scientific Report No. 7 under Contract No. AF 19(604)-4153 (Jan. 1962). Air Force Cambridge Research Laboratory document number AFCRL-62-129.

[2] G. Dresselhaus, A. Kip, and C. Kittel, Phys. Rev. 98, 368 (1955).

[3] R. Dexter, H. Zeiger, and B. Lax, Phys. Rev. 104, 637 (1956).

[4] W. Spitzer and H. Fan, Phys. Rev. 106, 882 (1957).

[5] B. Lax and G. Wright, Phys. Rev. Letters 4, 16 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ment of the frequency of the reflectivity minimum in the far infrared. Through a straightforward differentiation of the reflectivity equation, utilizing the dispersion relations, and the evaluation of this derivative at the minimum, a relationship is obtained which can be solved for m_c using an iterative procedure.

FREE-CARRIER DISPERSION

In the far infrared, at wavelengths longer than that of the intrinsic absorption edge, electromagnetic radiation interacts with the free carriers of the material and undergoes dispersion. The theory of this free-carrier dispersion is developed in detail in a recent book by Moss.[6] He has shown that both classical and quantum mechanical analyses lead to the following relations for the real refractive index n and the absorption index (extinction coefficient) k of free carriers of one type only (i.e., holes or electrons):

$$n^2-k^2-\epsilon_\infty=-\frac{Ne^2}{m\epsilon_0}\frac{\tau^2}{1+\omega^2\tau^2} \tag{2}$$

and

$$2nk\omega=\frac{Ne^2}{m\epsilon_0}\frac{\tau}{1+\omega^2\tau^2}, \tag{3}$$

where ϵ_∞ is the relative dielectric constant of the medium in the absence of any contribution from free carriers at very high frequencies, N is the free-carrier concentration, ϵ_0 is the permittivity of free space, m is the effective mass of the free carriers, and ω is the angular frequency of the infrared radiation.

Several points should be noted with regard to Eqs. (2) and (3). It has been shown for the energy-band model in which ellipsoidal surfaces of constant energy lie along rotation axes of the cubic point group, that the effective mass which enters in these expressions is the conductivity effective mass, m_c.[7] In this case, m_c is defined by the relation $1/m_c=\frac{1}{3}(1/m_l+2/m_t)$, where m_l is the value of the effective-mass tensor associated with the longitudinal axis of the ellipsoidal energy surface and m_t is that associated with each of the two transverse axes of the surface. Secondly, we note that the relaxation time τ is assumed to be independent of frequency and that the total frequency dependence is exhibited explicitly by the quantity ω. This assumption is not restrictive since the relaxation time is determined by the temperature of the crystal, its energy-band structure, and the density of imperfections in the lattice. Finally, it should be noted that the relaxation time τ is necessarily assumed to be energy-independent. This follows from the fact that the terms containing τ have not been averaged, $\langle\ \rangle$, with respect to the distribution function. This last assumption requires additional consideration. For example, we wish to establish the conditions under which τ may be replaced by $\langle\tau\rangle$ in Eqs. (2) and (3). The details of the analysis for establishing these conditions are presented in Appendix A. However, in summary if the relaxation time is assumed to be proportional to a power of the free-carrier kinetic energy $\tau=\tau_0E^s$, it is possible to compare the exact terms $\langle\tau^2/(1+\omega^2\tau^2)\rangle$ and $\langle\tau/(1+\omega^2\tau^2)\rangle$ with the more convenient substitute terms $\langle\tau\rangle^2/(1+\omega^2\langle\tau\rangle^2)$ and $\langle\tau\rangle/(1+\omega^2\langle\tau\rangle^2)$. Such a comparison shows that the latter terms can be substituted for the former under the following conditions with no introduction of error:

(i) when τ is independent of energy, $s=0$, as in the case of neutral impurity scattering,

(ii) when there is total degeneracy and the kinetic energy of all carriers may be considered to be the Fermi energy E_f.

The substitution can be effected under the following conditions with an introduction of error $<13\%$:

(iii) when $s=-\frac{1}{2}$, as in the case of scattering arising from acoustical modes and from dislocations, for all degrees of degeneracy where $\omega\tau\geqq1$ while the degree of degeneracy required increases from $\eta\leqq-4$ to $\eta=0$ as $\omega\tau$ decreases below unity,

(iv) when $s=\frac{1}{2}$, as in the case of optical-mode scattering above the Debye temperature, for all degrees of degeneracy and for all values of $\omega\tau$, and

(v) when $s=\frac{3}{2}$, as in the case of scattering by ionized impurities, only under the condition of total degeneracy. These conditions are not considered to be overly restrictive but, in fact, to be rather generally satisfied. This consideration follows particularly in view of the fact that in order to observe the free-carrier reflectivity minimum at convenient wavelengths ($\approx30\,\mu$), it is often necessary to utilize free-carrier concentrations which result in at least partial, if not total, degeneracy. For these reasons, in our further consideration of Eqs. (2) and (3), we will assume that one of the conditions (i)–(v) is fulfilled and that τ may be replaced by $\langle\tau\rangle$ with an introduction of error $<13\%$.

At this point, it must be recognized that the total dispersion of incident radiation in the far infrared arises from interactions with the free carriers and with the lattice as well. This latter dispersion is important in polar crystals. The complete dispersion relations, which include the contributions of both the free carriers and the lattice of a polar crystal AB, may be expressed as follows[8,9]:

$$n^2-k^2=\epsilon_\infty-\frac{Ne^2}{m_c\epsilon_0\omega^2}\frac{\omega^2\langle\tau\rangle^2}{(1+\omega^2\langle\tau\rangle^2)}+\left[\frac{\epsilon_\infty+2}{3}\right]^2\frac{N_tQ^2}{M\epsilon_0}\left[\frac{(\omega_t^2-\omega^2)\tau_t^2}{\omega^2+(\omega_t^2-\omega^2)^2\tau_t^2}\right] \tag{4}$$

[6] T. Moss, *Optical Properties of Semiconductors—a Semiconductor Monograph* (Butterworths Scientific Publications, London, 1959), Chap. 2, p. 29.

[7] Reference 6, p. 31.

[8] Reference 6, Chaps. 2 and 16.

[9] B. Szigeti, Trans. Faraday Soc. 45, 155 (1949).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and

$$2nk=\frac{Ne^2}{m_c\epsilon_0\omega^2}\frac{\omega\langle\tau\rangle}{(1+\omega^2\langle\tau\rangle^2)}+\left[\frac{\epsilon_\infty+2}{3}\right]^2\frac{N_lQ^2}{M\epsilon_0}\left[\frac{\omega\tau_l}{\omega^2+(\omega_t^2-\omega^2)^2\tau_l^2}\right]. \quad (5)$$

In both Eqs. (4) and (5) the last term is the contribution made by the ions of the crystal. The parameters in the other terms have been defined following Eqs. (2) and (3). The new parameters are as follows: N_l is the concentration of atom pairs AB; Q is an effective charge defined by $Q=SZe$, where e is the electronic charge, Z is the valence of the compound, and S is a measure of the extent to which the ions of charge Ze are capable of being polarized; M is the reduced atomic mass $1/M=1/M_A+1/M_B$; ω_t is the angular frequency of the transverse lattice vibrations (Restrahlen); and τ_l is a relaxation time related to the damping of the lattice vibrations.

The relative contribution of the lattice to the total dispersion varies from one semiconductor to another. Moreover, it even varies for a single semiconductor depending upon the free-carrier concentration. Consequently, this contribution must be evaluated for each individual case in order to determine if its effect is significant in the over-all dispersion. It is clear, however, that in several instances the lattice polarization may be neglected when analyzing the dispersion in the far infrared. For example, lattice polarization does not enter in connection with the elemental semiconductors, germanium and silicon. Moreover, a combination of several conditions, such as, a small value for ϵ_∞, a large reduced mass M, a slight polarizability ($S\ll1$), and a large free-carrier concentration ($N\gtrsim10^{18}$ carriers/cm^3), can serve to make the effect of the lattice polarization completely negligible.

Because various combinations of these conditions are quite often satisfied, the total dispersion in the far infrared can oftentimes be considered to be described by the simpler Eqs. (2) and (3). In the subsequent analysis we will make this assumption, namely, that the lattice contribution to the total dispersion is negligible in the frequency range of the reflectivity minimum arising from free-carrier absorption.

REFLECTIVITY MINIMUM

The reflectivity of a material for normally incident radiation upon the surface of an infinitely thick sample is given by[10]

$$R=[(n-1)^2+k^2]/[(n+1)^2+k^2], \quad (6)$$

where n and k are described by Eqs. (2) and (3). This equation has been differentiated with respect to the variable Ω after rewriting the dispersion relations in the form:

$$n^2-k^2=\epsilon_\infty-B/(1+\Omega^2) \quad (7)$$

and

$$nk=B/2\Omega(1+\Omega^2), \quad (8)$$

where

$$B=Ne^2\langle\tau\rangle^2/m_c\epsilon_0 \quad \text{and} \quad \Omega=\omega\langle\tau\rangle$$

and assuming that $\langle\tau\rangle$ is not a function of frequency. The condition of zero slope at the reflectivity minimum then yields an equation for the effective mass in terms of the frequency of the reflectivity minimum and other measurable parameters. The equation that results is the following:

$$m_c^{*3}-\frac{3\epsilon_\infty-1}{4\epsilon_\infty(\epsilon_\infty-1)}\frac{(1/\Omega^2)+5+8\Omega^2}{(1+3\Omega^2)}Cm_c^{*2}+\frac{3\epsilon_\infty-2}{2\epsilon_\infty(\epsilon_\infty-1)^2}\times\frac{(1+2\Omega^2)}{(1+3\Omega^2)}C^2m_c^*+\frac{C^3}{4\epsilon_\infty(\epsilon_\infty-1)^2(1+3\Omega^2)}=0, \quad (9)$$

where $m_c^*=m_c/m_0$, m_0 is the free-electron mass, $C=Ne^2/m_0\epsilon_0\omega_0^2$, $\Omega=\omega_0\langle\tau\rangle$, and ω_0 is the angular frequency of the reflectivity minimum. [The derivation of this equation is outlined in Appendix B. It is based entirely on Eqs. (6)–(8), and therefore depends on their applicability to a particular material system for its own validity.]

This equation is quite general in that no approximations are made with respect to the magnitude of Ω nor with respect to the magnitude of ϵ_∞. [However, it is true that Ω must staisfy the conditions under which the substitution of $\langle\tau\rangle$ into Eqs. (2) and (3) was effected.]

We conclude that there are very broad conditions for which Eq. (9) represents an exact relationship, or very nearly so, between the conductivity effective mass and the frequency of minimum reflectivity and other measurable parameters. Equation (9) can therefore be used for the determination of m_c. This can be illustrated by considering the reflection curve for n-type germanium reported by Spitzer and Fan (S-F) and which is reproduced here as Fig. 1. The sample investigated by S-F had a carrier concentration of 3.9×10^{18} electrons/cm^3, $\mu_h=520$ cm^2/V-sec,[11] $\epsilon_\infty=16$ and a reflectivity minimum at $\lambda_0\approx23\ \mu$. By making an initial estimate for the conductivity effective mass, say $m_c^*=0.12$ as determined by cyclotron resonance, one can obtain an initial estimate for Ω. Using this value of Ω and the values for N, ϵ_∞, and ω_0 listed above, Eq. (9) can be solved for m_c^*. This enables one to calculate a revised estimate for Ω and the procedure is repeated until the solution of Eq. (9) yields the same value for m_c^* as

[10] Reference 6, p. 6.

[11] It is recognized that we are here employing a Hall mobility when actually the drift mobility is called for. However, according to the S-F report, the Hall coefficient was measured at 77°K where the Fermi level is $\approx4.5\ kT$ above the conduction band edge for this free-carrier concentration. It is reasonable to conclude that the mobility deduced from $R\sigma$ is very nearly the drift mobility.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

before. In actual fact, because Ω^2 is large with respect to unity ($\Omega^2=13$) and therefore Eq. (9) is almost independent of Ω, this procedure converges very quickly. The value obtained for the conductivity effective mass using this procedure for the above example is $m_c^*=0.145$. This compares very well with the value $m_c^*=0.15$ calculated by S-F using their data from both reflection and transmission measurements.

As an additional check on the theory, for example, to show that the free-carrier dispersion can completely account for the observed reflectivity, a theoretical reflection curve can be generated using the effective mass calculated from Eq. (9). When this is done, one finds that the reflectivity curve, reported by S-F for the n-type sample of germanium considered above, is reproduced as exactly as it is possible to determine the values from their report. (The values calculated to achieve this fit are shown in Fig. 1.) Moreover, such a precise fit indicates that we are apparently justified in the substitution of $\langle\tau\rangle$ into Eqs. (2) and (3) and that in this case at least such a substitution produces negligible error. [This result should be compared with Fig. 4a (Appendix A) for $\Omega=3.6$, $\eta=-2$, and $s=-\frac{1}{2}$.]

Before concluding this discussion, it is of interest to consider the effect of various simplifying assumptions upon Eq. (9). For example, the plasma frequency is oftentimes obtained approximately by assuming that $n=1$ [see Eq. (6)], that k is negligibly small compared with unity, and that Ω is very much greater than unity. Under such conditions Eq. (7) can be solved directly to give the result

$$m_c=Ne^2/\omega_0^2\epsilon_0(\epsilon_\infty-1). \quad (10)$$

This expression is sometimes even further simplified by neglecting unity with respect to ϵ_∞ which yields

$$m_c=Ne^2/\omega_0^2\epsilon_0\epsilon_\infty. \quad (11)$$

Now, if the conductivity effective mass is calculated, for the n-type sample of germanium considered here, by means of Eqs. (10) and (11), the resulting values are $m_c^*=0.123$ and 0.115, respectively. These values differ significantly from the value $m_c^*=0.15$ obtained by Spitzer and Fan. This difference serves as justification for the derivation of the more exact Eq. (9).

Finally, it should be noted that Eq. (9) can be very well approximated by a quadratic equation since the constant term is oftentimes negligibly small.

$$m_c^{*2}-\frac{3\epsilon_\infty-1}{4\epsilon_\infty(\epsilon_\infty-1)}\frac{(1/\Omega^2)+5+8\Omega^2}{(1+3\Omega^2)}Cm_c^*+\frac{3\epsilon_\infty-2}{2\epsilon_\infty(\epsilon_\infty-1)^2}\frac{(1+2\Omega^2)}{(1+3\Omega^2)}C^2=0. \quad (12)$$

This equation can be solved directly, with the result that there are apparently two solutions. In actual fact, there is only one allowable solution. The root with the plus sign must be chosen, because it is the root which in the limit of $\Omega^2\gg1$ reduces to the known result expressed in Eq. (10).

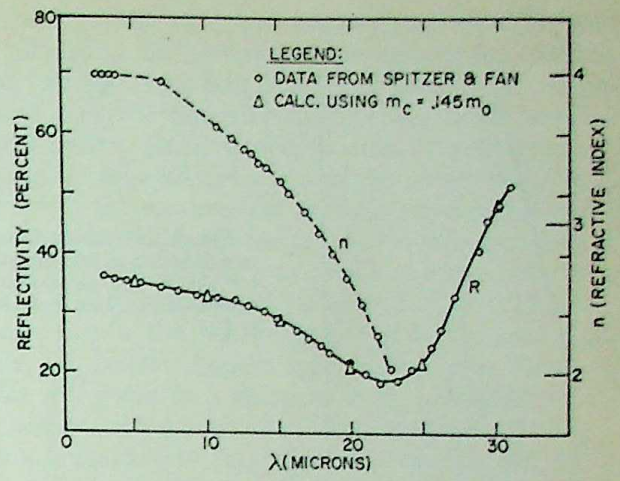


FIG. 1. Reflectivity and refractive index as functions of wavelength for n-type germanium having 3.9×10^{18} cm^{-3} as reported in Ref. 4. In addition, theoretical values for the reflectivity, having been calculated using an effective mass $m_c=0.145\ m_0$, are shown.

COMPARISON WITH OTHER TECHNIQUES

Cyclotron resonance techniques for the determination of the components of the conductivity effective mass have yielded a very great insight into the band structure of germanium and silicon (Ref. 12 gives a good review of this development). However, the condition for the observation of this resonance phenomenon is $\tau\omega_c>1$, where ω_c is the natural cyclotron frequency of rotation of a moving electron in a magnetic field about an axis parallel to the field, and classically this frequency is given by $\omega_c=eB/m_c$. Consequently, experiments in the microwave region, using reasonable magnetic field intensities, require that τ be 10^{-11} sec or longer. This condition is only fulfilled for high-purity semiconductors at low temperatures.

This limitation severely restricts the usefulness of cyclotron resonance for the determination of m_c. For example, some semiconductors have not as yet been prepared in a sufficiently pure form to enable cyclotron resonance measurements to be made. In this group, for example, is the very interesting family of semiconductors PbS, PbSe, and PbTe. Furthermore, there is reason to believe that the value of the conductivity effective mass in some semiconductors may be temperature-dependent. Evidence of this possibility has been reported in connection with investigations of both germanium[13] and lead sulfide.[14] However, because of the stringent condi-

[12] T. H. Geballe, *Semiconductors an American Chemical Society Monograph Series*, edited by N. B. Hannay (Reinhold Publishing Corporation, New York, 1959), Chap. 8, pp. 321–6.

[13] G. MacFarlane, F. McLean, J. Quarrington, and V. Roberts, Phys. Rev. 108, 1377 (1957).

[14] I. Smirnov, B. Moizhes, and E. Nensberg, Fiz. Tverd. Tela 2, 1992 (1960) [English transl.: Soviet Phys.—Solid State 2, 1793 (1961)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

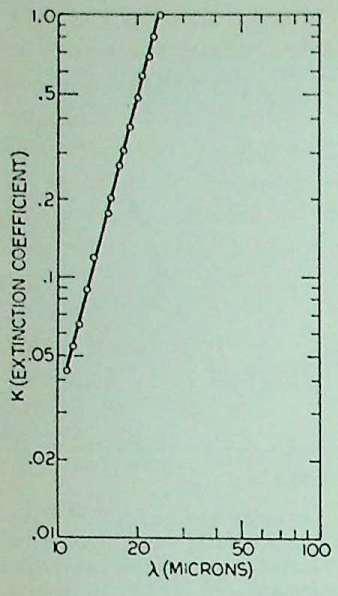


FIG. 2. Absorption index as a function of wavelength for the n-type germanium sample of Fig. 1 as reported in Ref. 4.

tion $\tau\omega_c > 1$, cyclotron resonance measurements are generally restricted to liquid-helium temperatures and so they cannot be utilized for the determination of the temperature dependence of m_c.

More recently, techniques have been developed which have utilized measurements in the far infrared. As mentioned earlier (S-F) obtained values of m_c from a combination of reflection and transmission measurements,[4] and Lax and Wright (L-W) obtained their values from the reflectivity of polarized radiation for several values of magnetic field intensities[5] (magnetoplasma effect). In both of their analyses (S-F and L-W), the simplification of neglecting the contribution of the lattice to the total dispersion has been inferred.

In connection with the S-F procedure, one notes that if the condition $(\omega\tau)^2 \gg 1$ is fulfilled, the dispersion relation of Eq. (2) is independent of τ. Under this constraint, the conditions placed upon the energy dependence of τ, as noted earlier, do not apply. That is, under the constraint that $(\omega\tau)^2 \gg 1$, the relaxation time may have an arbitrary energy dependence without having any effect on the resultant dispersion relationship, which is

$$n^2 - k^2 - \epsilon_\infty \approx Ne^2/m_c\epsilon_0\omega^2. \tag{13}$$

In addition to the simplified dispersion relation of Eq. (13) and the reflectivity relationship expressed in Eq. (6), the S-F procedure utilizes the fact that the transmission through a sample of material of thickness x in the absence of interference fringes and for $k^2 \ll n^2$ is given by[15]

$$T = (1-R)^2 e^{-\alpha x}/(1-R^2 e^{-2\alpha x}), \tag{14}$$

where α is the linear absorption coefficient which is related to the absorption index k by the relation

$$\alpha = 4\pi k/\lambda. \tag{15}$$

[15] Reference 6, pp. 13–14.

The linear absorption coefficient α equals the reciprocal of the distance within which the incident radiation is attenuated by a factor $(\exp)^{-1}$. The procedure outlined by S-F for the determination of m_c is readily deduced from Eqs. (6), (13)–(15).

This method for obtaining the conductivity effective mass of free carriers is relatively simple compared with cyclotron resonance techniques. It does require that $(\omega\tau)^2 \gg 1$. However, since the frequencies are of the order of 10^{13}–10^{14} rad/sec, τ may be of the order of 10^{-13} sec compared with 10^{-11} sec for cyclotron resonance. Consequently, measurements can be made on less pure materials and at higher temperatures. Furthermore, it may or may not require transmission measurements. Under certain conditions where k increases very rapidly with wavelength (see Fig. 2), it is possible to consider the reflection region where $k^2 \ll (n-1)^2$ and thus obtain m_c directly from Eqs. (6) and (13). However, one must beware that going to wavelengths much smaller than λ_0 means that $n^2 \to \epsilon_\infty$ and small errors in the absolute reflectivity can cause very large errors in m^* calculated from Eqs. (6) and (13). Consequently, in general, the S-F procedure requires transmission measurements which in turn require the preparation of very thin samples. For their transmission measurements, S-F used samples whose thickness was of the order of 10 μ. Such thicknesses represent a few thousand atomic layers and are very often difficult to achieve without producing extensive damage to the remaining material. For example, it has been possible to achieve such thicknesses in lead telluride only with very careful grinding and polishing techniques and then only very small surface areas ($\approx 600 \times 50\ \mu$) were obtained so that special optical techniques must be utilized.[16]

The other procedure mentioned above, the measurement of the magnetoplasma reflection as proposed by L-W,[5] eliminates the necessity of making transmission

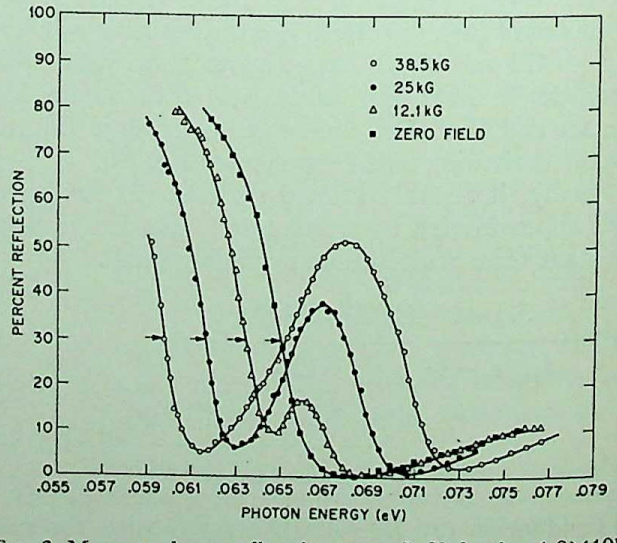


FIG. 3. Magnetoplasma effect in n-type InSb having 1.8×10^{18} cm^{-3} as reported in Ref. 5.

[16] W. W. Scanlon, Phys. Chem. Solids 8, 423 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

measurements. This method utilizes the fact that, in the presence of properly polarized radiation, a magnetic field shifts the minimum in the reflectivity to both sides of the zero-field minimum. This results in the appearance of two minima. The displacement of these minima to either side of the zero-field minimum is directly proportional to the cyclotron resonance frequency $\omega_c = eB/m_c$. In their investigation, L-W have determined the shift as a function of magnetic field strength and have determined m_c from the slope of this dependence. Their data for the magnetoplasma effect in n-type indium antimonide are shown in Fig. 3.

This procedure, even though it offers the advantage of eliminating transmission measurements, does contain several inherent drawbacks. For example, magnetic field strengths of 25-kG or less would produce a shift in the minimum of approximately $\pm\frac{1}{2}\,\mu$ for the germanium sample of Fig. 1. Consequently, for zero-field reflection minima which are shallow and which have a width of a couple of microns, as for example the reflectivity minimum observed by S-F for germanium at room temperature in Fig. 1, the actual shift of the minimum is very difficult to measure. In fact, in their original article Lax and Wright found that the positions of the shifted minima were sufficiently indistinct so that it was expedient to determine the shift from that of the reflection edge on the long-wavelength side of the reflection minimum at an isoreflection point. This shift is that indicated by the arrows in Fig. 3. However, this latter method for obtaining the shift of the reflection minimum is also subject to certain disadvantages. First of all, it is not clear that the edge should shift the same amount as the minimum since the latter shift was determined assuming that losses are negligible. This is certainly not the case in the region of the reflection edge on the long-wavelength side of the reflection minimum because in this region we have the condition $k>n$. Furthermore, because k increases very rapidly in this region of the reflectivity, the depth of penetration of the incident radiation becomes increasingly shallow. For example, from Eq. (15) we see that the penetration depth $(1/\alpha)$ for $\lambda=30\,\mu$ and $k\approx 2$ is only of the order of $1\,\mu$. Consequently, the surface preparation becomes increasingly important in this region of the reflectivity curve. Surface damage can therefore distort the reflectivity in this wavelength range. Such surface effects make it even more difficult to determine the relationship of the observed shift in the reflection edge to the cyclotron resonance frequency.

CONCLUSIONS

The relationship between m_c and ω_0 presented in Eq. (9) offers some unique advantages in the determination of m_c. It has rather widespread applicability since it is valid for totally degenerate statistics and for almost all degrees of degeneracy when $|s|\leq\frac{1}{2}$, where s is the power of the energy dependence of τ. Equation (9) is not limited by any assumption that $\omega\tau>1$. It is much more applicable than the simplified relation of Eq. (11). It eliminates the necessity of determining the transmission through very thin samples of material on the one hand and the necessity of applying large magnetic fields perpendicular to the direction of propagation of polarized infrared radiation on the other. What is required is the measurement of the infrared reflectivity in order to determine the frequency of minimum reflection. It is not even necessary to obtain absolute reflectivity data. Relative values can serve equally well. That is, a percentage of the incident and/or of the reflected radiation can be lost through experimental error. Although this will result in a decrease in the resolution of the minimum, if the minimum is nevertheless discernible, it is still possible to determine m_c from Eq. (9). Moreover, this procedure is not overly sensitive to the effects of surface damage on the long-wavelength side of the reflectivity minimum. It has been noted above that the surface preparation becomes increasingly important in this region because of the decreasing depth of penetration into the sample. A damaged sample surface has the tendency of dispersing the radiation thereby reducing the reflectivity in this long wavelength region and of presenting a $\langle\tau\rangle$ which is different from that of the bulk material. The reflectivity in the region of the minimum is also subject to distortion from this mechanism only less so than in the region of the absorption edge because of the increased depth of penetration of the incident radiation.

Consequently, m_c can be obtained using Eq. (9) over the large temperature range in which the reflectivity minimum is measurable. The free-carrier concentration of the semiconductor material of interest is limited only by the requirement of single-carrier conduction, by the requirement that the lattice contribution to the dispersion be relatively unimportant, and by the requirement that the reflectivity apparatus be capable of analyzing the reflection near the free-carrier dispersion minimum. However, these limitations on the applicability of this procedure are also common to the other procedures which utilize infrared radiation.

ACKNOWLEDGMENT

I am very greatly indebted to Professor Arthur C. Smith of the Electrical Engineering Department at the Massachusetts Institute of Technology for his advice on the matters presented here.

APPENDIX A

It is convenient to substitute terms expressed as functions of $\langle\tau\rangle$ into Eqs. (2) and (3). Such a substitution is only approximately equivalent to the exact terms. The nature of the approximation involved can be appreciated by examining the following ratios:

$$\frac{\langle\tau^2/(1+\omega^2\tau^2)\rangle}{\langle\tau\rangle^2/(1+\omega^2\langle\tau\rangle^2)} \quad \text{and} \quad \frac{\langle\tau/(1+\omega^2\tau^2)\rangle}{\langle\tau\rangle/(1+\omega^2\langle\tau\rangle^2)} \qquad (16, 17)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

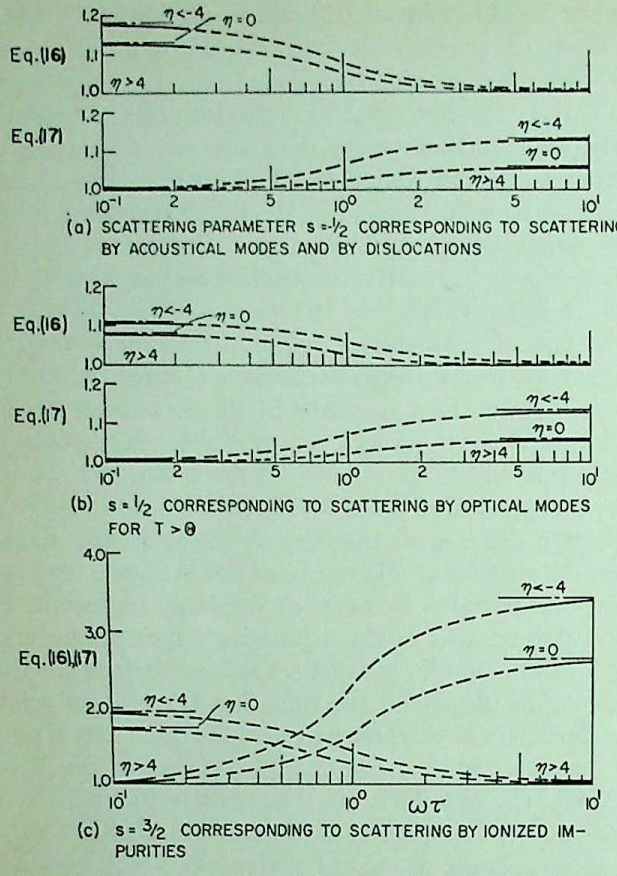


FIG. 4. Asymptotic limits of the ratios (16,17) plotted as functions of $\omega\tau$ for various values of the scattering parameter s and for three degrees of degeneracy $\eta<-4$, $=0$, >4.

under the conditions of $\omega\tau\gg1$ and $\omega\tau\ll1$. Under these conditions the above ratios approach the following limiting values:

$$\text{for } \omega\tau\gg1\text{: } 1 \text{ and } \langle\tau^{-1}\rangle\langle\tau\rangle$$
$$\text{for } \omega\tau\ll1\text{: } \langle\tau^{2}\rangle\langle\tau\rangle^{-2} \text{ and } 1.$$

Under these limiting conditions, it is apparent that for complete degeneracy these ratios are unity. In the limit of nondegeneracy these limiting values become

$$\text{for } \omega\tau\gg1\text{: } 1 \text{ and } \Gamma(5/2-s)\Gamma(5/2+s)[\Gamma(5/2)]^{-2}$$
$$\text{for } \omega\tau\ll1\text{: } \Gamma(5/2+2s)\Gamma(5/2)[\Gamma(5/2+s)]^{-2} \text{ and } 1$$

The asymptotic limits for various scattering mechanisms are shown in Fig. 4 for reduced Fermi potentials of $\eta<-4$, $=0$, and >4.

Appendix B

The condition $\partial R/\partial\Omega=0$ yields the equation

$$(n^2-1-k^2)n'+2nkk'=0, \tag{18}$$

where the prime signifies $\partial/\partial\Omega$. Solving for k and k' from Eq. (8) and substituting into (18) yields

$$(n^4-n^2-3\beta^2)n'+2\beta n\beta'=0, \tag{19}$$

where $\beta=B/[2\Omega(1+\Omega^2)]$. Combining Eqs. (7) and (8) yields

$$n^4=\epsilon n^2+\beta^2, \tag{20}$$

where $\epsilon=\epsilon_\infty-[B/(1+\Omega^2)]$, and upon substituting (20) into (19), we obtain

$$[(\epsilon-1)n^2-2\beta^2]n'+2\beta n\beta'=0. \tag{21}$$

By solving (20) for n^2 (choosing the positive root to keep n real), by differentiating this result with respect to Ω, and by substituting $(n^2)'/(2n)$ for n' in Eq. (21), we obtain

$$[(\epsilon-1)n^2-2\beta^2][-2n^2\Omega^3+\beta(1+3\Omega^2)]$$
$$+2n^2\beta(1+3\Omega^2)(2n^2-\epsilon)=0. \tag{22}$$

Equation (22) can be rewritten in the form

$$n^4=G(\Omega)n^2+H(\Omega)\beta^2, \tag{23}$$

where $G(\Omega)$ and $H(\Omega)$ are complicated functions of Ω. Equating (23) and (20) yields

$$n^2=\beta^2(H-1)/(\epsilon-G). \tag{24}$$

Dividing Eq. (20) by n^2 and substituting for n^2 from Eq. (24) yields the relationship

$$\beta^2=(\epsilon-G)(H\epsilon-G)/(H-1)^2. \tag{25}$$

After much simplification (25) reduces to Eq. (9) of the text.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 4A 18 MAY 1964

Pressure Effect on Resistivity of $Ga(As_{1-x}P_x)$

G. E. Fenner
General Electric Research Laboratory, Schenectady, New York

(Received 2 January 1964)

The change of resistance of n-type $Ga(As_{1-x}P_x)$ samples under hydrostatic pressure up to 15 000 atm was measured between 190 and 363°K. Large pressure-dependent resistance changes were observed for samples with $x \approx 0.35$. These observations are consistent with the presence of another conduction band edge of much lower mobility presumably along $k = (100)$ into which electrons are transferred when the relative energy separation between the two edges is lowered by pressure. The measured results are compared with calculations based on the model proposed by Ehrenreich. The agreement between theory and experiment is found to be good if certain parameters are varied in a systematic way. For lightly doped material the mobility ratio of electrons in the light mass band to those in the heavy mass band is found to be large and to decrease rapidly with increasing temperature. At higher carrier concentrations both the magnitude of the mobility ratio as well as its variation with temperature are substantially reduced. At 300°K the (000) and (100) minima are found to cross at $x \approx 0.44$ if we assume Vegard's law to be valid. The experimental data are best explained if it is assumed that the separation between the two sets of minima changes at a rate of $>1\times10^{-4}$ eV/°K.

INTRODUCTION

In the last few years considerable progress has been made in the understanding of compound semiconductors, in particular those consisting of elements from the III–V columns in the periodic table. A comprehensive review of the transport properties of a number of these has been given by Ehrenreich,[1] who also deduced the band structure of $Ga(As_{1-x}P_x)$ as a function of the composition variable x. The details have been exhaustively covered by that author in an earlier publication[2] and therefore we shall only review those points of his model that are of direct concern to us here. Unlike the majority of semiconductors investigated to date, $Ga(As_{1-x}P_x)$ displays two sets of conduction band edges which are close enough in energy over an appreciable portion of the composition range so that their combined influence on the electrical behavior of the semiconductor has to be considered. The separation in energy between the two sets of minima depends upon x and it vanishes for $x \approx 0.5$. For $x < 0.5$ the (000) minimum is lowest in energy, while for $x > 0.5$ the (100) valleys are believed to be lowest. Because of the large differences in the mobilities and density of states between them, the bulk properties of the material can be expected to change appreciably as the region $x = 0.5$ is approached. It is the purpose of this paper to report measurements made in that transition region under the application of hydrostatic pressure and to derive values for parameters that so far have been speculative or not known at all. The influence of hydrostatic pressure on the band structure of semiconductors has been discussed by Paul.[3] He also touches upon some of the discrepancies in the experimental results in GaAs, which have a direct bearing on the model of both that compound and that of $Ga(As_{1-x}P_x)$ itself.

Our approach in this paper is to assume the correctness of the model proposed in Ref. 2. We hope to determine several band-structure parameters by fitting our results to this model. As will become apparent shortly, the biggest difficulties in accomplishing our task arise from the lack of homogeneity in the crystals themselves.

EXPERIMENTAL DETAILS

The $Ga(As_{1-x}P_x)$ samples used for the measurements were obtained from various sources. The most homogeneous samples were cut from crystals grown by halogen-assisted vapor transport in a sealed capsule somewhat similar to the method reported by Pizzarello.[4] Others were grown from excess gallium solution and one was an epitaxially grown sample purchased from the Merck Company. All crystals were doped with Te. X-ray determination of the lattice constant in conjunction with Vegard's law was used to establish the composition most accurately. All the crystals grown by vapor transport were prepared to yield $x \approx 0.35$. Whenever possible the specimens were cut into rectangular bars about 10 mm long and 1 mm square. Ohmic contacts were made by alloying tin dots to the semiconductor. The pressure was generated in a cylindrical multiple wall pressure cell of conventional design.[5] The pressure generating piston was driven directly by the platen of a 50-ton press. At room temperature the measurements were carried out inside this cylinder, while at the other temperatures the samples were enclosed in a small pressure cell similar to the one reported by Fitchen.[6] The large pressure cylinder then was used as the pump and connected to the small cell by $\frac{1}{8}$-in. o.d., 0.024-in. i.d. hard-drawn stainless-steel tubing manufactured by the Harwood Engineering Company. The electrical connections were made by the use of

[1] H. Ehrenreich, J. Appl. Phys. 32, 2155 (1961).
[2] H. Ehrenreich, Phys. Rev. 120, 1951 (1960).
[3] W. Paul, J. Appl. Phys. 32, 2082 (1961).
[4] F. A. Pizzarello, J. Electrochem. Soc. 109, 226 (1962).
[5] C. A. Swenson in *Solid State Physics*, edited by F. Seitz and D. Turnbill (Academic Press Inc., New York, 1960), Vol. 11, p. 41; P. W. Bridgman, *The Physics of High Pressure* (G. Bell and Sons, London, 1949).
[6] D. B. Fitchen, Rev. Sci. Instr. 34, 673 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

0.062-in. swaged metallic sheathed thermocouple wire.[7] At temperatures below 300°K a mixture of n pentane and isopentane served as the pressure fluid while at temperatures above that n pentane alone was used. The pressure was measured by a seasoned manganin wire gauge[5] mounted inside the pressure cylinder. The gauge calibration was established using a recent determination of the freezing pressure of Hg at 0°C.[8] The accuracy of the pressure measurements is estimated to be better than 1%.

RESULTS

The application of hydrostatic pressure on Ga($As_{1-x}P_x$) increases the energy gap between the conduction band minimum at $k=(000)$ and the top of the valence band at the rate of about 1.1×10^{-5} eV/atm. The energy separation between the (100) minima and the valence band has a small negative pressure coefficient and therefore the net result is a decrease of the relative energy separation ΔE between the (000) and (100) conduction band valleys under pressure. As long as $\Delta E \gg kT$ the number of electrons transferred to the energetically higher lying minima per unit dilatation of the crystal is small. Once ΔE has been reduced to a few kT either by very high pressures in the case of [2]GaAs or by the proper choice of composition in Ga(AsP), the rate of transfer of electrons to the higher valleys increases rapidly. Because of the much lower electron mobility in the (100) valleys, the bulk resistance of the material increases many fold until all the current is carried by electrons in the (100) valleys. After that a further increase in pressure does not change the resistance appreciably. This general characteristic is demonstrated by the measurements shown in Figs. 1, 3, and 4.

As mentioned already our plan is to compare experiment and theory by choosing certain parameters of Ehrenreich's model until a best fit with the measured values is obtained. With the assumptions set down in the Appendix we find for the relative resistance at pressure P

$$\frac{\rho(P)}{\rho(0)}=\frac{b+\nu}{1+\nu}\frac{1+\nu e^{\alpha P/kT}}{b+\nu e^{\alpha P/kT}}, \qquad \text{(A1)}$$

where $b=\mu_1/\mu_2$ is the mobility ratio between valleys 1 and 2, $\nu=b_\rho e^{-\Delta E(0)/kT}=n_2(0)/n_1(0)$ the ratio of the number of electrons in valley 2 to the number in valley 1 at atmospheric pressure, and α the pressure coefficient of the energy separation between valleys 1 and 2. Valley 1 refers to the minimum at $k=(000)$ while 2 refers to the set of equivalent minima along $k=(100)$. In trying to fit Eq. (A1) to the experimental data, the value for the pressure coefficient α was accepted from earlier work. It was deduced directly from pressure measurements on [2]GaAs and probably can be accepted with some confidence.[9] Should a small correction for this value become necessary it would merely lead to a minor correction in b as seen by Eq. (A4) of the Appendix. The density of states ratio b_ρ is numerically quite uncertain[2] and may be subject to correction. Fortunately for most of our considerations b_ρ is absorbed into the factor ν and will only become important when the band structure at a particular value of x is discussed.

Figure 1 shows the results of measurements made on sample 21a-11, the crystal with the smallest carrier concentration of all our samples. This single-crystal specimen obtained from a crystal grown by vapor transport also had a small variance in composition and therefore we will base our conclusions largely on that particular sample. Other bars cut from adjacent parts of the same crystal gave results very close to 21a-11, so that our confidence seems justified. Except for the high-pressure region, i.e., the region after the bands have crossed, the fit is considered to be very good. The outstanding feature of the curves plotted in Fig. 1 is the behavior of $b=\mu_1/\mu_2$ as a function of temperature. It has a very large value at 200°K but decreases markedly as the temperature is increased. For the deviations between the expected and the actual measured values at

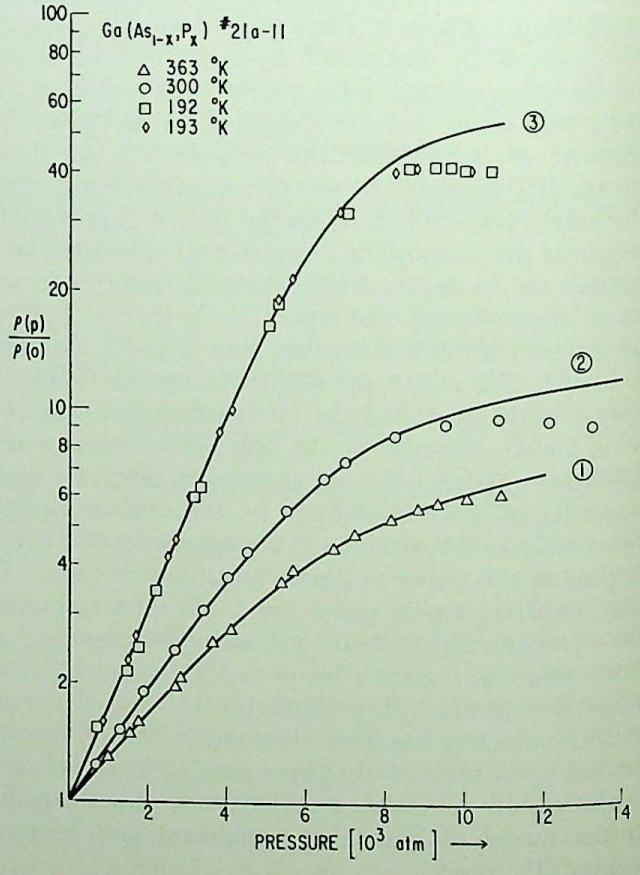


FIG. 1. Relative resistance versus pressure: Ga($As_{1-x}P_x$), No. 21 a-11, $x=0.359$. Curve 1: $\nu=7.2$, $b=55$; curve 2: $\nu=5.6$, $b=75$; curve 3: $\nu=4.8$, $b=320$.

[7] R. H. Cornish and A. L. Ruoff, Rev. Sci. Instr. 32, 639 (1961).
[8] D. H. Newhall, L. H. Abbot, and R. A. Dunn, American Society of Mechanical Engineers, New York Meeting, 1962 (unpublished).
[9] For GaP this coefficient has been found to be larger by about 25%. See Ref. 3.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

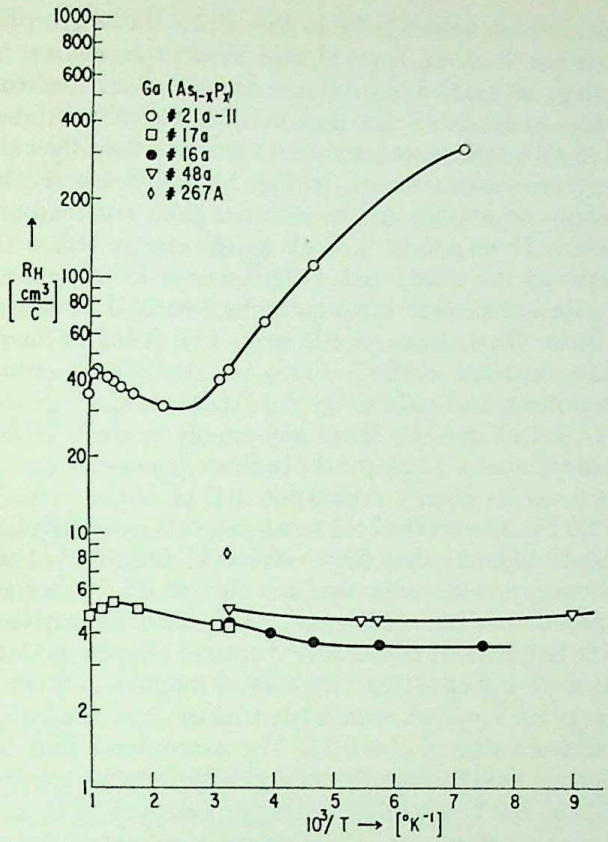


FIG. 2. Hall constant versus $1/T$ at atmospheric pressure.

high pressures, we offer the following explanation: The inspection of the Hall coefficient (Fig. 2) for this sample reveals that freeze out of carriers occurs below 400°K. This implies that an appreciable number of carriers is occupying one or more deep levels at the temperatures at which our experiments are performed. If we suppose that these levels have a pressure shift equal to that of the (000) minimum, then the number of electrons in that band will not change under pressure as long as the other minima are sufficiently far removed in energy. As the (000) minimum approaches the (100) minima, some of the carriers in the impurity levels can now be thermally excited into the (100) minima. Now our assumption that the total number of electrons in the conduction band remains constant is no longer valid. Because of the large mobility of electrons in the (000) minimum, this effect will not be noticeable until the latter valley is almost empty; i.e., until the bands have crossed. Then, as carriers are transferred from the bound impurity states to the low but finite mobility (100) band, the resistance decreases in agreement with the observations. The effect should become more pronounced as the temperature is decreased, because the ratio of electrons in the (000) minimum to the number in bound states is decreased even before the pressure is applied. For the samples that do not show the freeze out, this decrease in resistance does not occur, giving further support to the correctness of our reasoning.

Figure 3 shows the results of measurements made on sample 17a. This specimen is also a single crystal and are similar in composition to 21a-11. The major difference between the two crystals is the much larger carrier concentration of 17a. This difference shows up in the reduction of the mobility ratio b which can be traced to the increased influence of ionized impurity scattering in the heavier doped crystal. Because this scattering process is most important at low temperatures, it reduces the temperature dependence of b.

The results of measurements on additional samples are shown in Fig. 4. Specimen 48a is a polycrystalline sample which shows a pressure dependence similar to 17a. However, the factor ν which results in the best fit with the experimental data has a different temperature dependence for the two samples. It is probable that the polycrystalline sample 48a has a larger variation in composition which would explain the different behavior. A very pronounced effect of this variation can be seen in sample 267 A with $x=0.24\pm0.05$. This epitaxially grown single crystal shows such large variations in composition that no agreement between our model and experiment can be accomplished.

Finally, we have measured a sample with a composition of $x\approx0.61$ and find the change in resistance linear up to 10 000 atm at room temperature with $\Delta\rho/(\rho P) = -14.5\pm1\times10^{-6}$ atm^{-1}. This agrees in sign with the

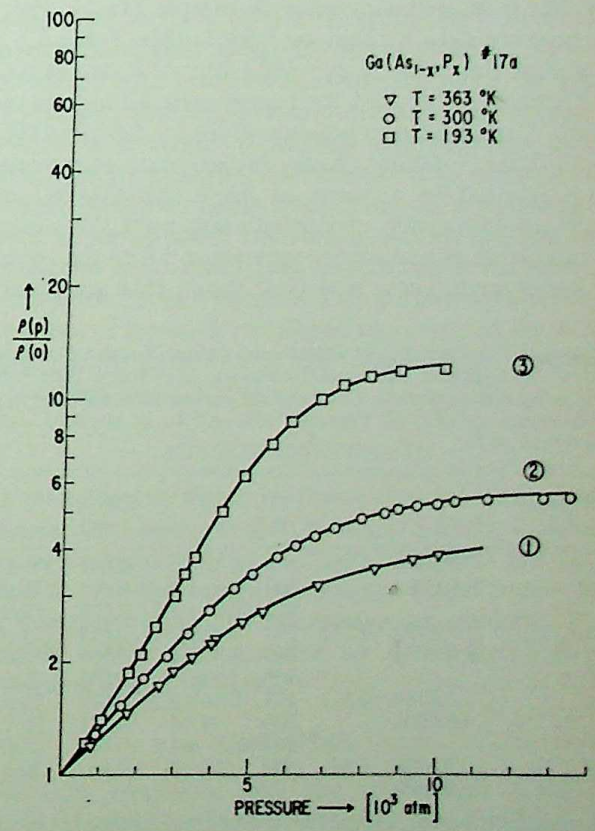


FIG. 3. Relative resistance versus pressure: $Ga(As_{1-x}P)$, No. 17a, $x=0.372$. Curve 1: $\nu=5.3$, $b=22$; curve 2: $\nu=3.3$, $b=22$; curve 3: $\nu=0.85$, $b=23$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

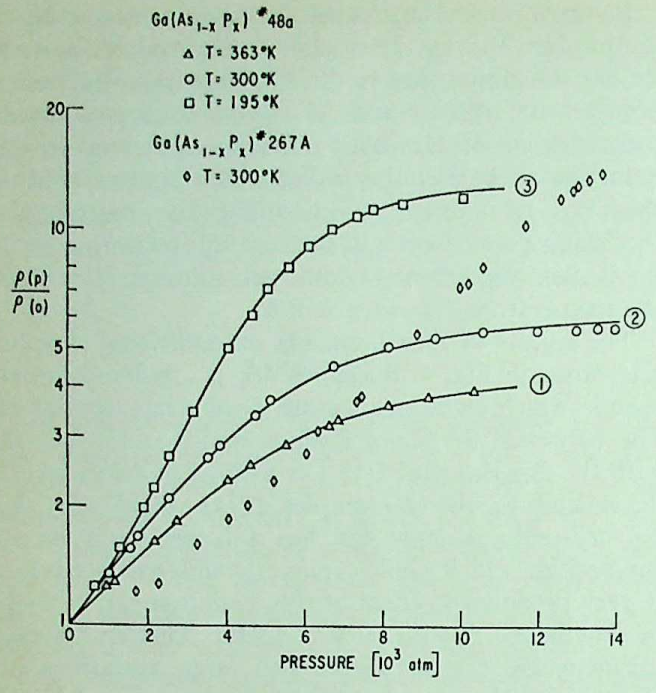


FIG. 4. Relative resistance versus pressure: $Ga(As_{1-x}P)$ No. 48a, $x=0.367$, and No. 267*A*, $x=0.24$. Curve 1: $\nu=6.1$, $b=25$; curve 2: $\nu=4$, $b=26$; curve 3: $\nu=0.85$, $b=23$.

measurements made on[10] silicon and[11] GaP, but is larger in magnitude than either one of these.

We shall now attempt to gain additional information from the measurements made on sample 21a-11. From the pressure data we find $\nu=5.6\pm1$. (See Table I.) If we accept the value of $b_\rho=70(+40, -20)$ determined by Ehrenreich,[2] we find the energy separation between valleys 1 and 2 at zero pressure given by $\Delta E(0)=0.065$ eV $(+0.017, -0.013)$. X-ray lattice-parameter evaluation resulted in a value of $a_0=5.5806\pm0.0002$ or, assuming the validity of Vegard's law, $x=0.359$. Linear extrapolation of our energy separation ΔE to the terminal compounds GaAs and GaP shows that we cannot match both values given in Ref. 2 simultaneously. The pressure data of Howard and Paul[12] rule against an energy separation between the direct and indirect band edges in GaAs of less than 0.3 eV. Ehrenreich found 0.36 eV as the most consistent value based on Hall and pressure measurements. In Fig. 5 we have plotted the energy separation ΔE as deduced from our measurements. If we accept 0.36 eV as the energy separation between the (000) and (100) minima for GaAs, the separation for GaP turns out to be 0.46 eV if we assume a linear dependence of ΔE on x. This is 0.1 eV larger than reported earlier.[1] From an analysis of optical absorption and reflectivity data Zallen and Paul[13] find 2.75–2.8 eV for the direct gap energy in GaP. With a value of about 2.2 eV for the indirect gap energy E_G this leads to an energy separation ΔE of about -0.5 to -0.6 eV. Photothreshold measurements by Spitzer and Mead[14] indicate that $\Delta E\approx-0.45$ eV. Engeler,[15] of this laboratory, concludes that a value of 0.5 eV for the separation of the (100) and (000) band is consistent with Spitzer and co-workers'[16] optical absorption data. Figure 5 indicates the two sets of minima to cross in energy at $x_c\approx0.44$, somewhat smaller than the earlier deduced value of $x_c\approx0.53$. The assumption that ΔE depends linearly on x is undoubtedly only an approximation, but the deviations are not expected to be large enough to affect our determination of x_c materially.

TABLE I. Experimentally determined values of ν and b as function of temperature for several samples. α_0 = lattice parameter from x-ray measurements. Composition parameter x deduced from α_0 assuming validity of Vegard's law. $\Delta E(0)$ is nominal value based on $b_\rho=70$.

Sample No.	α_0 [Å]	x	T [°K]	ν	$\Delta E(0)$ [eV]	$b=\mu_1/\mu_2$
21a-11	5.5806	0.359	193	4.4 ±1	0.045	300±50
	±0.0002		300	5.6 ±1	0.065	75±15
			363	7.2 ±1.2	0.075	55±15
17a	5.5780	0.372	193	0.85±0.2	0.073	23±3
	±0.0001		300	3.3 ±1	0.079	22±3
			363	5±1	0.087	21±3
48a	5.5789	0.367	193	0.85±0.2	0.073	23±3
	±0.0001		300	4±1	0.074	26±5
			363	6.1 ±1.1	0.08	25±4
NH 1 No. 4	5.5738	0.39	300	10±2	0.05	30±7
	±0.0004					

The main difficulties in the interpretation of our results show up when we try to determine the effect temperature has on the energy separation ΔE. Again accepting the value $b_\rho=70$ independent of temperature, we can find ΔE from ν. This is done in Table I for three samples discussed earlier. In order to arrive at the probable error of these determinations, curves calculated by varying both b and ν were compared with the actual measurements. The limit of agreement was defined as that point at which the deviations of the calculated curves from the measured have passed outside the experimental tolerances. The discrepancies in the temperature coefficient of ΔE for the three samples are larger than the experimental error. In view of the difficulties arising from crystal inhomogeneities, as exemplied by results from sample 267 A, this is not too surprising. It must be concluded, though, that $\partial\Delta E/\partial T=(\partial/\partial T)(E_{100}-E_{000})$ is larger than 1×10^{-4} eV/°K.

While the changes of resistance under pressure can readily be understood in terms of the adopted model, the interpretation of the Hall data derived from the same samples is not at all certain. The Hall coefficient for the two-band model is given by

$$R_H=[(1+\nu)/en][(b^2+\nu)/(b+\nu)^2],$$

[10] C. S. Smith, Phys. Rev. **94**, 42 (1954).
[11] A. Sagar and R. C. Miller, J. Appl. Phys. **32**, 2073 (1961).
[12] W. E. Howard and W. Paul, quoted in Ref. 3.
[13] R. Zallen and W. Paul (to be published).
[14] W. G. Spitzer and C. A. Mead (to be published).
[15] W. E. Engeler (private communication).
[16] W. G. Spitzer, M. Gershenzon, C. J. Frosh, and D. F. Gibbs, Phys. Chem. Solids **11**, 339 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

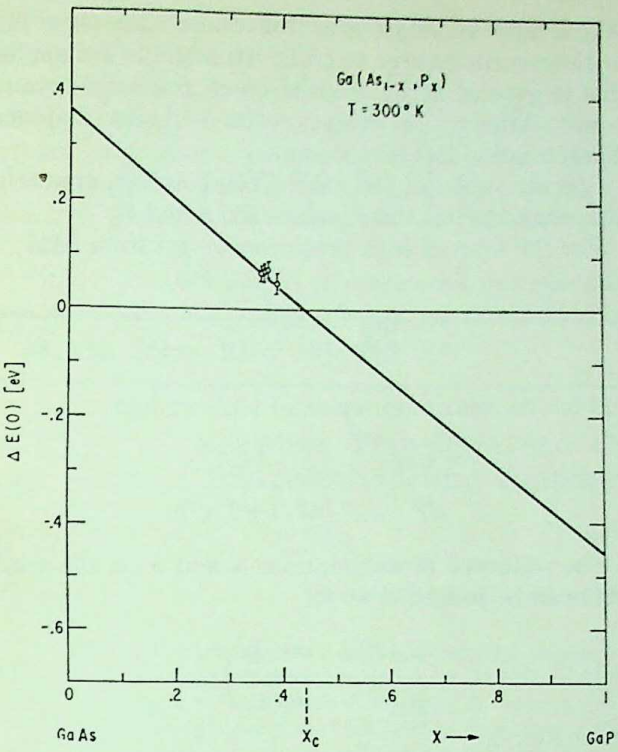


FIG. 5. Energy separation $\Delta E(0)$ versus composition at atmospheric pressure and 300°K. The uncertainty in the experimental points is due to the experimental errors as well as the large uncertainty in the knowledge of $b\rho$.

where scattering factors of the order of 1 have been set equal to 1. Under the condition $n_1+n_2=n=$constant, R_H has a maximum value of $(1+b)^2/4b$ when $n_2/n_1=b$. The temperature at which this maximum occurs increases with b and ΔE. To be consistent with the parameters deduced from the pressure experiments, the Hall coefficient should have increased appreciably at room temperature over its value at lower temperatures assuming ΔE to be temperature-independent. This, however, is contrary to the experimental evidence. Furthermore, the maximum of R_H occurs at much higher temperatures than expected from the energy separation of the (000) and (100) valleys at room temperature.

For the lightly doped sample 21a-11, the rapid freeze out below 400°K invalidates our assumption that the total number of electrons in the conduction band remains constant. This fact and the rapid decrease of b with increasing temperature could be advanced as a reason for the poor agreement with the measured Hall coefficient. Both arguments are not supported by the measurements made on the rest of the samples and, therefore, we are forced to look for other possible explanations. As mentioned earlier, the pressure measurements at various temperatures are best explained if we assume ΔE to have a positive temperature coefficient in excess of 1×10^{-4} eV/°K. An increase of the energy separation ΔE at a rate of about 2×10^{-4} eV/°K and a decreasing b would suffice to keep R_H approximately constant to higher temperatures, except that this is contrary to the experience with the other group IV and III–V semiconductors as well as the terminal compounds themselves. It should be noted, however, that very little is known about transport properties of ternary compounds and some of the extrapolations might well be in error.

CONCLUSIONS

The close agreement between the measured values and the calculations based on Ehrenreich's model of the $Ga(As_{1-x}P)$ conduction band must be considered strong evidence for the correctness of the essential points of that model. While we have not directly proved that the energetically higher lying minima are indeed along the ⟨100⟩ directions, the self-consistency of our results gives strong support to this identification.

The large mobility ratio ($b>50$) deduced from our room temperature measurements on the lightly doped sample is significant, since this is necessary for the interpretation of the large pressure-dependent resistance changes in GaAs observed by Howard and Paul.[12] The rapid decrease of the mobility ratio b with increasing temperature could explain the smallness of b derived from high-temperature Hall data. Extended to GaAs, it suggests that this is the reason for the apparent difference in $b=10$ reported by Aukerman and Willardson[17] and the b based on room-temperature pressure measurements. But in view of the discrepancies between our pressure and Hall data, this interpretation must be considered with caution. The composition x_c at which the (000) and (100) minima are equal in energy is smaller than earlier estimates, excepting the recent determination of Spitzer and Mead.[14] The extrapolation of their results to GaAs leads to an energy separation which is too small. Their results can be brought into agreement with ours if we assume some variation in composition in their samples. This variation would have the effect of lowering the apparent energy of the (100) band edge for a given value of x explaining the difference in their results.

ACKNOWLEDGMENTS

I would like to thank H. Ehrenreich, W. Paul, R. O. Carlson, W. Engeler and R. N. Hall for many informative and helpful discussions; S. J. Silverman and T. J. Soltys for the Hall measurements and G. J. Charney for his assistance with the pressure measurements. The crystals were kindly supplied by R. F. Reihl, S. J. Silverman, and R. N. Hall.

APPENDIX

We are assuming that the conduction band of Ga(As,P) exhibits two sets of minima, one minimum at

[17] L. W. Aukerman and R. K. Willardson, J. Appl. Phys. 31, 939 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$k=[000]$ (valley 1) and a set of minima along $k=[100]$ in the Brillouin zone (valley 2). The separation in energy between the minima at the different points in the Brillouin zone is supposed to be a linear function of the composition of the alloy as shown in Fig. 5. For the purpose of the analysis we make these further simplifying assumptions:

1. Boltzman statistics apply.
2. Intervalley scattering is neglected.
3. The total number of carriers in the conduction band is constant.
4. The mobility ratio $b=\mu_1/\mu_2$ is independent of pressure. (Subscript 1 refers to the low-mass band at $k=[000]$ throughout this analysis. Furthermore, symbols used are those of Ehrenreich.[2])
5. $b_\rho=(m_\rho/m_1)^{3/2}=70$ is also pressure- and temperature-independent; $m_\rho=$density of state mass in the valleys at $k=[100]$.
6. The energy separation between different minima is $\Delta E=\Delta E(0)-\alpha P$ where $\alpha=1.1\times10^{-5}$ eV/atm.

With these assumptions we find for the relative resistance at pressure P,

$$\frac{\rho(p)}{\rho(0)}=\frac{b+\nu}{1+\nu}\frac{1+\nu e^{\alpha P/kT}}{b+\nu e^{\alpha P/kT}}, \tag{A1}$$

where

$$\nu=\frac{n_2(0)}{n_1(0)}=\left(\frac{m_\rho^{3/2}}{m_1}\right)e^{-\Delta E(0)/kT}.$$

The logarithmic derivative of the expression (A1) is given by

$$\frac{d\ln\rho}{dP}=\left(1-\frac{1}{b}\right)\frac{\alpha}{kT}\frac{\nu e^{\alpha P/kT}}{[1+(\nu/b)e^{\alpha P/kT}](1+\nu e^{\alpha P/kT})}. \tag{A2}$$

Since ν can directly be converted into composition by the use of Fig. 5, this last equation gives essentially a universal curve in which composition and pressure can be interchanged. In principle, it should therefore be possible to plot the experimental values of the logarithm of resistance versus pressure, differentiate this curve and fit the resulting curve to (A2). Attempts to accomplish that in general were not satisfactory, primarily because b was different in every crystal and the graphical differentiation lacked accuracy.

Specific values of (A1) and (A2), however, are useful in finding the unknown parameters ν and b.

For the limit of high pressures, we get from (A1)

$$\lim_{P\to\infty}\frac{\rho(P)}{\rho(0)}=\frac{b+\nu}{\nu+1}, \tag{A3}$$

and for the maximum value of (A2) we find

$$\left.\frac{d\ln\rho}{dP}\right|_{\max}=\frac{\alpha}{kT}\frac{1-1/b}{1+2/\sqrt{b}}. \tag{A4}$$

The influence of assumptions 4 and 5 on the result (A1) can be judged if we let

$$\frac{\mu_1(P)}{\mu_1(0)}=\left(\frac{m_1(0)}{m_1(P)}\right)^n. \tag{A5}$$

From Ref. 2,

$$\frac{m_1(P)}{m_1(0)}=\frac{E_G(P)}{E_G(0)}\frac{1+(1/2)E_G(0)[E_G(0)+\Delta]^{-1}}{1+(1/2)E_0(P)[E_G(P)+\Delta]^{-1}}, \tag{A6}$$

where $E_G(0)$ and $E_G(P)$ are the band-gap energy at atmospheric pressure and at pressure P respectively and Δ the spin-orbit splitting energy. If polar scattering is the dominant scattering mechanism, n in equation (A5) equals $\frac{3}{2}$ and (A1) can be corrected by replacing the one in the numerator by the right hand side of (A6) raised to the $\frac{3}{2}$ power. Since $\Delta\ll E_G$ (A6) is approximately

$$1+\frac{1}{E_G(0)}\left(\frac{\partial E_G}{\partial P}\right)P.$$

This factor has to be considered only for values of ν and P where the electron transfer to the higher band edge is insignificant.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, Vol. 134, No. 5A 1 JUNE 1964

Resistive Transition and Current Density Characteristics in Superconducting Niobium Containing Dissolved Gases*

WARREN DESORBO
General Electric Research Laboratory, Schenectady, New York
(Received 5 March 1963; revised manuscript received 4 February 1964)

Oxygen and nitrogen as interstitial solutes in niobium cause the metal in transverse field to exhibit anomalous resistance minima or critical current maxima (peak effect). The anomaly appears at H_{c2}, the upper critical field of these Type II superconductors. The resistance in the mixed state is decreased by exceeding the solubility limit, by cold working and by strain aging at 170°C, masking out the anomaly. Rapidly quenching from the solution temperature decreases the anomaly. The results suggest that the anomaly is strongest at some intermediate stage of the process of segregation and precipitation at dislocations. Resistance in the mixed state is increased by introducing more solute, by quenching, and by etching. Structural changes brought about by cold working, strain aging, and rapid quenching responsible for large changes in $J_c(H)$ or $R(H)$ in the mixed state also cause changes in the region $H_{c2}<H<H_{c3}$. In this latter region, field orientation effects and observed H_{c3}/H_{c2} ratios are roughly consistent with the concept of a superconducting surface sheath as recently proposed by Saint-James and de Gennes. Pure niobium, however, tends to exhibit consistently higher ratios.

I. INTRODUCTION

SUPERCONDUCTING transition metals in Group V of the periodic table have been noted for their anomalous behavior which is usually attributed to the presence of chemical and/or structural impurities. An important metal of this group is niobium, which serves as a base metal in the fabrication of many high-field, high-current-carrying superconductors. The influence of substitutional[1] and interstitial[2] solutes on some superconducting properties of noibium has been recently summarized. Other investigations have also appeared, revealing the importance of small amounts of interstitial impurities on the substructures of the metal (e.g., see Ref. 3).

Unusual superconducting behavior occurring in niobium has been reported in both the resistance-field[4] and critical current-field characteristics.[5] Recently, some striking resistance decreases following an initial rise were observed[6] during the superconducting transition of the metal: (a) when the current I through the wire, and temperature T were kept constant and a transverse magnetic field was increased continuously; or (b) when the current and field were held constant and T varied. The field at which the resistance minimum takes place was found to be independent of J, occurring at a field comparable to the upper critical field, H_{c2} of a Type II superconductor.[7,8] It is pointed out in this report that this anomalous behavior is readily observed in niobium containing small amounts of interstitial solutes such as oxygen or nitrogen but is not observed in outgassed niobium. The results are compared with the magnetic behavior of the solutions which have already been described[2] as Type II superconductors. The influence of solute concentration, field orientation, cold working, strain aging, quenching, and etching on the resistance-field transition and on the current

* The contents of this paper were presented in part before the American Physical Society meetings, Washington, D. C., 22–25 April 1963 and Philadelphia, Pennsylvania, 23–26 March 1964.

[1] W. DeSorbo, Phys. Rev. 130, 2177 (1963).

[2] W. DeSorbo, Phys. Rev. 132, 107 (1963).

[3] J. O. Stiegler, C. K. H. Dubose, R. E. Reed, Sr., and C. J. McHargue, Acta. Met. 11, 851 (1963).

[4] T. G. Berlincourt, Phys. Rev. 114, 969 (1959).

[5] M. A. R. LeBlanc and W. A. Little, in *Proceedings of the Seventh International Conference on Low-Temperature Physics*, edited by G. M. Graham and A. C. Hollis-Hallet (University of Toronto Press, Toronto, 1961), p. 362.

[6] S. H. Autler, E. S. Rosenblum, and K. H. Gooen, Phys. Rev. Letters 9, 489 (1962).

[7] A. A. Abrikosov, Zh. Eksperim. i Teor. Fiz. 32, 1442 (1959). [English transl. Soviet Phys.—JETP 5, 1174 (1957)].

[8] B. B. Goodman, IBM J. Res. Develop. 6, 63 (1962).

Copyright © 1964 by The American Physical Society.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

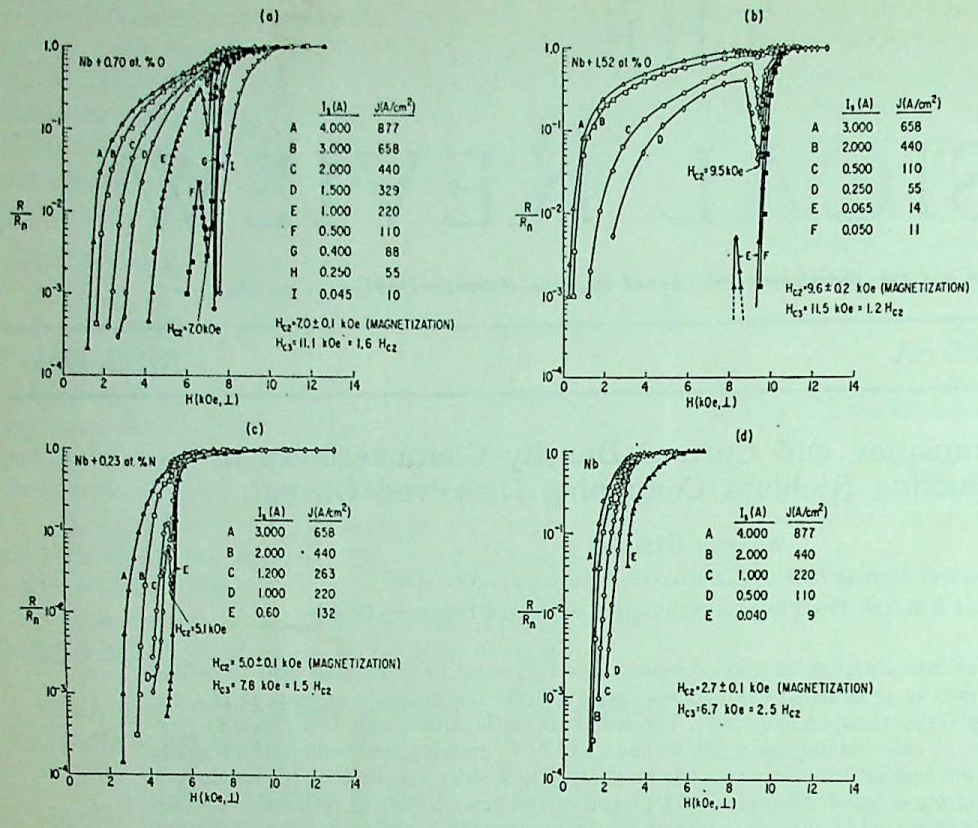


FIG. 1. The resistive characteristics R/R_n as a function of current density and applied transverse field at 4.2°K for wire (0.030-in. diam) specimens of (a) Nb+0.70 at.%O (T_c=8.7$_8$°K); (b) Nb+1.52 at.%O (T_c=8.0$_4$°K); (c) Nb +0.23 at.% N (T_c=9.20°K); and Nb (T_c=9.4$_6$°K). H_{c3} values are evaluated at $J\approx10$ A/cm².

density-field characteristics both in the mixed state below H_{c2} and in the region $H_{c2}\lesssim H\lesssim H_{c3}$ are summarized. The dependence of H_{c3}—the field at which the specimen is completely normal—on these factors is cited.

II. EXPERIMENTAL

The niobium wire specimens (0.030-in. diam) were made from an electron-beam melted stock obtained from the Stauffer-Temescal Company, Richmond, California, hereafter referred to as Nb(S-T) (see Ref. 2). Rods, about 0.250-in. diam were machined from the ingot then swaged and drawn to wire form. The resistance ratio, $R_{298°K}/R_{10°K}$, of the wire "as received" was about 110. The quantitative addition of the interstitial solute was carried out by a procedure developed earlier by Powers and Doyle[9]. The interstitial solute was introduced at temperatures between 1000–1100°C. The specimen was then homogenized by heating at about 1200–1300°C for approximately 1 h. Prior to the introduction of the gas, the specimen had been annealed and outgassed at about 1800–2000°C for several hours at a residual pressure less than 1×10^{-7} mm Hg. After this treatment and prior to the introduction of the interstitial, the resistance ratio of the wire was about 500. An analysis of a similarly prepared specimen reported by Seybolt[10] gave oxygen content 6±3 ppm and nitrogen content 5±3 ppm.

The resistance transition as a function of field at constant current density was measured at 4.2°K potentiometrically with a noise level less than 0.1 μV. Referred to in terms of R/R_n, where R is the resistance in the transition and R_n, the normal resistance, the noise level is less than 1×10^{-3} except for data obtained with low current densities ($J<125$ A/cm²) for outgassed niobium. The critical current of the wire specimen in a given field was also measured potentiometrically and the value taken was that which produced approximately 0.1 μV across the potential leads. The fields were produced by a small niobium wire (0.004-in. diam) wound iron-core magnet (0.040-in. slot) designed and built by H. R. Hart, Jr., following an original design due to Autler.[11] Field direction was in all cases transverse to the wire and to current flow. Field orientation for ribbon samples, either perpendicular ($H\perp$w.s.) or parallel ($H\|$w.s.) to the wide surface were obtained by mounting the ribbon in a form permanently fixed between the pole pieces.

The magnetic induction measurements were carried out by a technique used earlier by Bean.[12] A typical

[9] R. W. Powers and Margaret V. Doyle, J. Appl. Phys. **28**, 255 (1957); J. Metals **9**, 1285 (1957).

[10] A. U. Seybolt (private communication).

[11] S. H. Autler, D. B. Montgomery, and G. Ajootian, Quarterly Progress Report on Solid State Research, Lincoln Laboratory, MIT, July 1960, p. 63 (unpublished).

[12] C. P. Bean, Marget V. Doyle, and A. G. Pincus, Phys. Rev. Letters **9**, 93 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sample prepared for these measurements consisted of 7–10 wire segments, approximately 1.2 cm long, bundled together into a cylindrical form in a paper "straw." The ends of each wire were etched (50% HF, 50% HNO_3) to remove any strains. The sample cylinder was immersed in liquid helium at zero field. Magnetization measurements were made at 4.2°K by flipping it in and out of a copper coil (about 2000 turns) coaxial with an external steady field. The coil was connected to a flux meter whose deflection was proportional to the magnetization of the specimen. The magnetic field measurements were accurate to ±10 Oe; while $-4\pi M$ values were accurate to ±5 Oe.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Influence of Oxygen and Nitrogen in Niobium

1. Resistive Transition at Constant Current Density

The resistive transition R/R_n as a function of transverse field H and current density J at 4.2°K for four wire specimens of niobium containing (a) 0.70 at.% oxygen, (b) 1.52 at.% oxygen, (c) 0.23 at.% nitrogen, and (d) niobium outgassed and annealed is shown in Fig. 1. The resistance decrease in the transition, after an initial rise, is apparent for each of the interstitial solid solutions of niobium containing oxygen or nitrogen at current densities below a maximum value J_{max} and above a minimum value J_{min}. The drop in resistance is most pronounced at current densities near J_{min}. Increasing the interstitial concentration in the homogeneous solid solution tends to decrease both J_{min} and J_{max}. For example, for specimen Nb+0.70 at.% oxygen, $J_{max}\sim 1000$ A/cm² and $J_{min}\sim 90$ A/cm²; while for specimen Nb+1.52 at.% oxygen, $J_{max}\sim 800$ A/cm² and $J_{min}\sim 12$ A/cm². For a given specimen, the field corresponding to the maximum decrease in R/R_n, after an initial rise, appears to be almost independent of J and coincides with H_{c2}, the upper field required for $-4\pi M=0$ (compare Fig. 1 and Fig. 2).

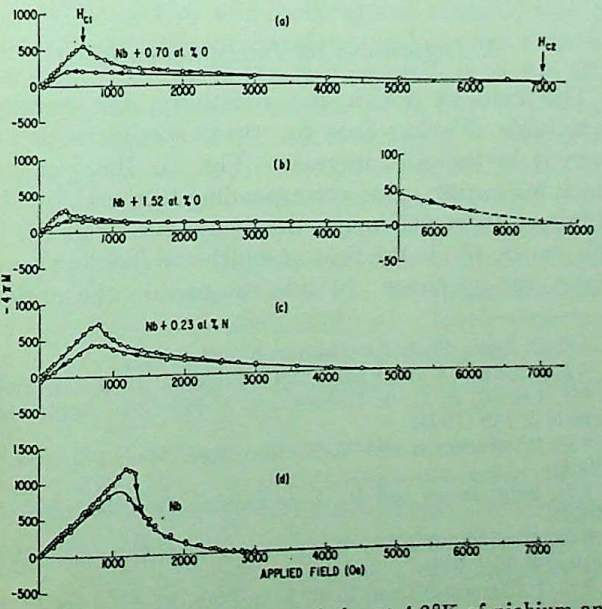
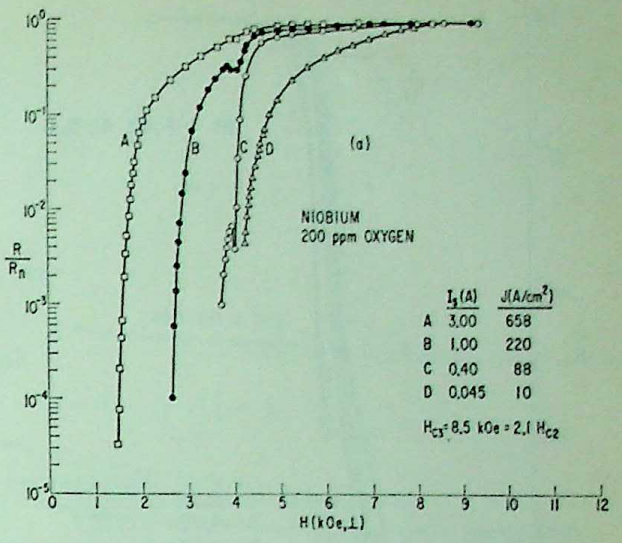


FIG. 2. Magnetization characteristics at 4.2°K of niobium and niobium containing interstitial solutes of oxygen and nitrogen (0.030-in. diam wires).



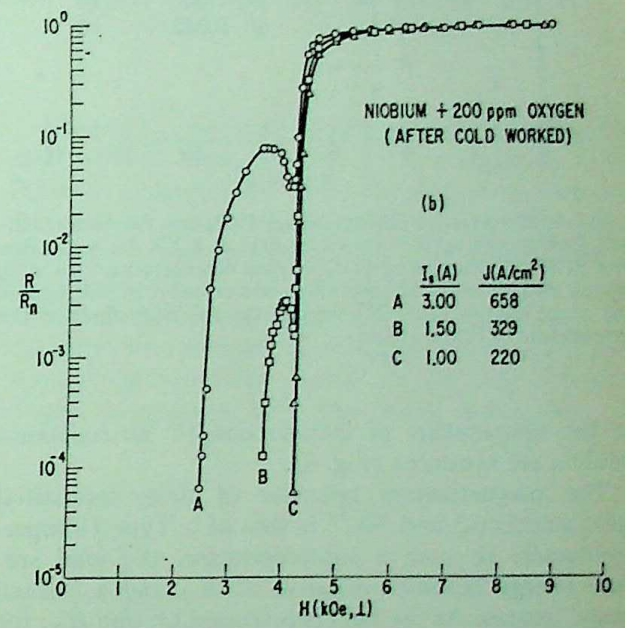


FIG. 3. The resistive transition R/R_n as a function of applied field at 4.2°K for a niobium wire containing 200 ppm oxygen (a) before cold worked diam=0.030 in.), H_{c2} evaluated at $J=10$ A/cm² and (b) after cold worked by hammering (diam ≈0.027 in.). $H_{c2}=8.0$ kOe$=1.9\ H_{c2}$ at $J=200$ A/cm².

For niobium outgassed and annealed, no apparent discontinuities in the transition are seen even at low current densities [Fig. 1(d)]. At these small J values the noise level in R/R_n is greater than 1×10^{-3}. However, the anomaly can still be observed in niobium containing as little as 200 ppm oxygen [Fig. 3(a) and (b)]. When the oxygen is beyond the solubility limit

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

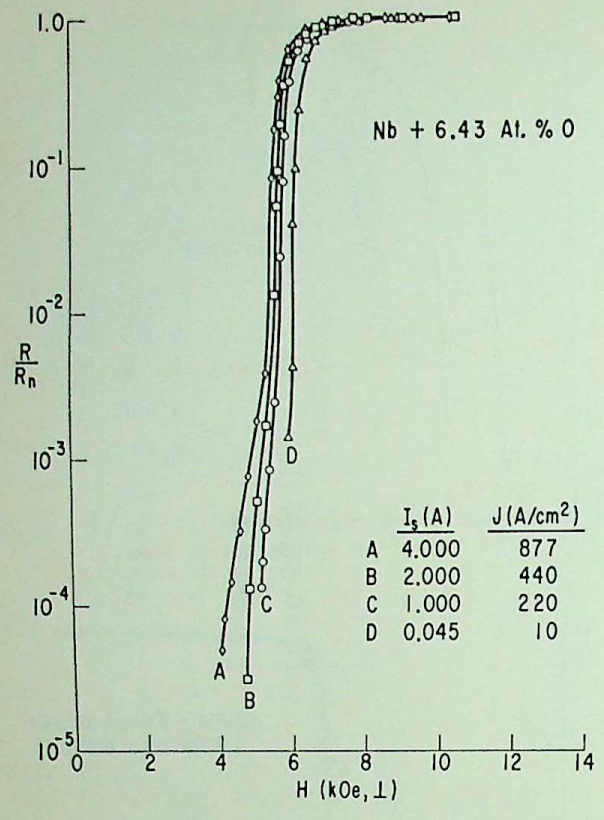


FIG. 4. The resistive characteristics R/R_n as a function of current density and applied tranverse field at 4.2°K for a niobium wire (0.030-in. diam) containing oxygen concentration [6.4 at.% oxygen plus some nitride contamination revealed by micrograph and x-ray studies (Ref. 2)] beyond the solubility limit at the temperature of introduction.

at the temperature of introduction,[13,14] no resistance minima are apparent (Fig. 4).

The magnetization behavior of these interstitial solid solutions,[2] and Nb,[15] is that of a Type II superconductor. In such a superconductor, the total surface energy becomes negative at H_{c1} and a "mixed state" ensues. As the field is increased beyond H_{c1}, the flux penetration increases until at H_{c2} the transformation to the normal state is achieved. The model is based on the material being homogeneous and defect-free, exhibiting reversible magnetization characteristics. It is believed that the resistance $R(H)$ in the "mixed state" is associated with flux motion or creep of the kind discussed by Kim, Hempstead, and Strnad,[16] i.e., resistance is determined from the rate at which flux filaments are moving. Based on this model, increasing the interstitial concentration seems to increase the flux motion for a given J since R increases [compare Fig. 1(a) and 1(b)]. When the solubility limit of the interstitial is exceeded, no resistance appears below H_{c2} within the limits of the measurements (compare Fig. 4 and Fig. 9, Ref. 2). In the "mixed state" the interaction of defects, inhomogeneities, or precipitates with flux filaments controls the hysteresis of the type described by Bean[17,18] and the equivalent critical current.[16,19]

Anderson and Kim[20] have suggested a possible explanation of the anomalous resistance drop after an initial rise based on the flux-creep model. As H_{c2} is approached, the lattice flux lines are expected to become more rigid, slowing down the rate of flux motion. Oxygen and nitrogen present in small concentrations would tend then to favor this process.

Above H_{c2} the variations in resistivity appear at smaller current densities dependent on the measuring currents. R approaches R_n at H_{c3}, the field at which the material is completely normal. There are two possible explanations for the variation of the $R(H)$ curve in this region. One is that recently proposed by Saint-James and de Gennes[21] envisioning a superconducting sheath along parts of the sample which disappears at H_{c3}. Recently, experimental evidence[22–24] has appeared showing that this remanent superconductivity is basically a surface effect in agreement with this point of view. Some supporting evidence for this is also presented in this work (see below). An alternative explanation may be that above H_{c2} a real resistance occurs; and that the shape of this portion of the $R(H)$ curve manifests the destruction of the final traces of superconductivity due to lattice irregularities such as dislocations, grain boundaries, and inhomogeneities having a higher critical field than the matrix. The connected regions would play an important role near and just above H_{c2}, while the discontinuous superconducting regions would be responsible for the shape of the curve at the top of the transition.

2. *Influence of the Interstitial on H′*

The field at which the resistivity first becomes detectable H' decreases for these specimens as the current in the wire increases (Fig. 5). It appears to reach a limiting value corresponding to the value H_{c1}, obtained from the magnetization curve (Fig. 2). In this curve, H_{c1} is the field at which the flux first penetrates the specimen. In this comparison the applied

[13] A. U. Seybolt, Trans. AIME 200, 774 (1954).
[14] E. Gebhardt and R. Rothenbacher, Z. Metalk. 54, 623 (1963).
[15] T. F. Stromberg and C. A. Swenson, Phys. Rev. Letters 9, 370 (1962).
[16] Y. B. Kim, C. F. Hempstead, and A. R. Strnad, Phys. Rev. 131, 2486 (1963).
[17] C. P. Bean, Phys. Rev. Letters 8, 250 (1962).
[18] J. Silcox and R. W. Rollins, Appl. Phys. Letters 2, 231 (1963).
[19] J. Friedel, P. G. de Gennes, and J. Matricon, Appl. Phys. Letters 2, 119 (1963).
[20] P. W. Anderson and Y. B. Kim, Rev. Mod. Phys. 36, 39 (1964).
[21] D. Saint-James and P. G. de Gennes, Phys. Letters 7, 306 (1964).
[22] G. Bon Mardion, B. B. Goodman, and A. Lacaze, Phys. Letters 8, 15 (1964).
[23] C. F. Hempstead and Y. B. Kim, Phys. Rev. Letters 12, 145 (1964).
[24] W. J. Tomasch and A. S. Joseph, Phys. Rev. Letters 12, 148 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

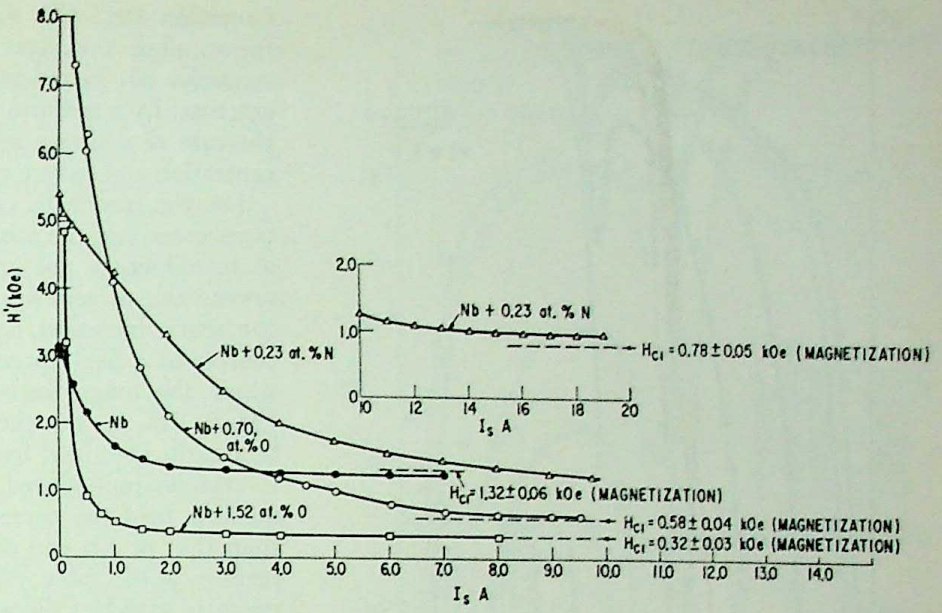


FIG. 5. The transverse field H' required for the initial appearance of resistance ($R/R_n \lesssim 1\times10^{-3}$) at 4.2°K as a function of specimen current I_s (amperes). The values of H_{c1} are evaluated from the magnetization curve.

field value does not contain the small correction due to the magnetic field generated by the current in the wire. H' decreases as the interstitial solute concentration increases. The resistive transition occurring between H' and the field at which the resistance minimum occurs, H_{c2}, for a given J becomes broader as the solute concentration increases [compare Figs. 1(a) and 1(b)]. More quantitative comparisons should be done at the same reduced temperature since oxygen present interstitially lowers the transition temperature.[2]

3. Critical Current Density

The critical current density J_c as a function of transverse field, of wire specimens of annealed niobium and annealed niobium containing various concentrations of oxygen is shown in Fig. 6. J_c extends to higher fields as the oxygen concentration increases. This is no longer true when the solubility limit is exceeded. A comparison of the curves with their respective magnetization data (Fig. 2) clearly defines three portions: (1) below H_{c1}, a sharp decrease in $J_c(H)$ takes place analogous to the behavior in a "soft" superconductor; (2) the "mixed state" region between H_{c1} and H_{c2}, where the anomalous behavior in the threshold-field curve (peak effect) related to the resistance minimum discussed above is evident in the niobium containing 0.70 at.% oxygen and (3) a region above H_{c2}.

Below H_{c1}, one assumes that the transport current J is carried in a thin shell (London penetration depth). Between H_{c1} and H_{c2}, it is believed that the transport current is controlled by the interaction of defects and flux filaments and that the net flow of current is dependent on the concentration gradient of flux filaments resulting from this interaction. J_c is determined as soon as the motion of the flux filaments is detected.[16,19]

Above H_{c2}, evidence cited below seems to indicate that current could be carried by a superconducting layer as suggested by Saint-James and de Gennes.[21] An alternative explanation could be that starting at H_{c2}, the continuous supercurrent flux filaments in the main body of the matrix no longer exist and the current carrying mechanism changes to that associated with extended defects consisting of higher critical field superconducting structural defects or physical filaments. Similar suggestions have been made earlier (e.g., see

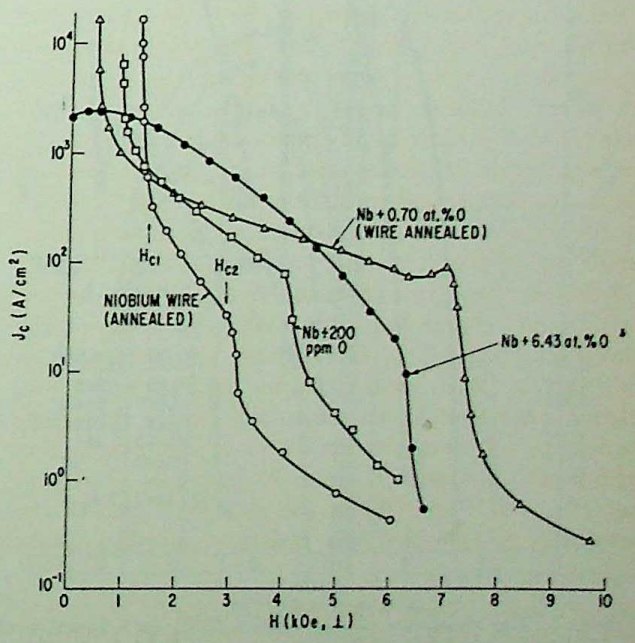
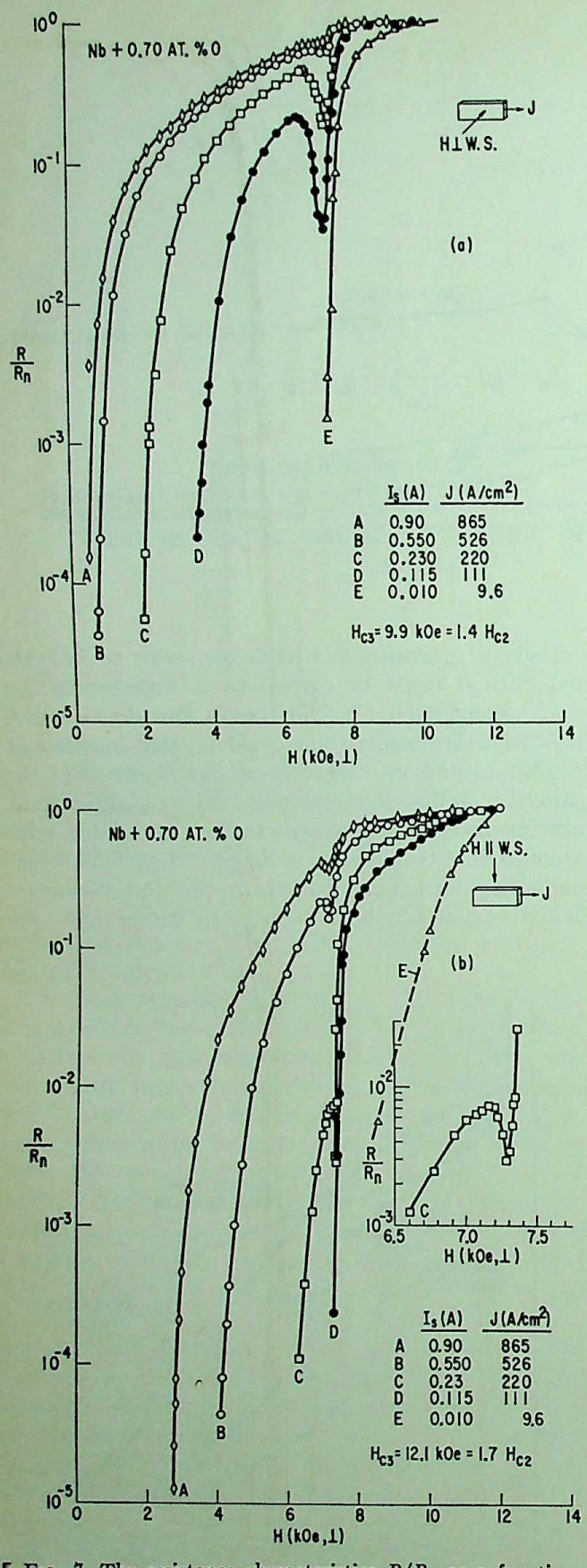


FIG. 6. Threshold current density-field behavior of some annealed niobium wires (0.030-in. diam) containing various amounts of oxygen. Annealed and outgassed niobium is also shown for comparison ($T=4.2$°K).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri



F FIG. 7. The resistance characteristics R/R_n as a function of applied transverse field at 4.2°K for an annealed ribbon of niobium (0.032×0.007 in.) containing 0.70 at.% oxygen (a) with field perpendicular to the wide side ($H\perp$w.s.); (b) with field parallel to the wide side ($H\|$w.s.). H_{c3} values are evaluated at $J\simeq 10$ A/cm².

discussion Ref. 25). $J_c(H)$ drops off rapidly in this region since the superconducting paths would, presumably, not be continuous nor necessarily be characterized by a uniform critical field. One would expect the rate of decrease of $J_c(H)$ to depend on the concentration and nature of the defects.

For the specimens containing oxygen in concentrations exceeding the solubility limit at the temperature of introduction, the critical current has a relatively lower value at zero field, remaining almost constant as the field is increased, decreases monotonically and then sharply at a field approximately corresponding to that where the magnetization becomes zero ($B=H$) (see Fig. 9, Ref. 2). For this specimen H_{c1} and H_{c2} cannot be clearly identified from the magnetic behavior since it exhibits pronounced magnetic hysteresis. At intermediate field the current density is somewhat higher than that of Nb and dilute Nb-O solutions indicating perhaps some more effective interaction of the precipitates with flux filaments in the Nb matrix.

B. Influence of Field Orientation

1. Resistive Transition at Constant Current Density

a. Nb+0.70 at.% Oxygen-Annealed Ribbon. Figures 7(a) and 7(b) show the resistance transition of an annealed niobium ribbon containing 0.70 at.% oxygen for $H\perp$w.s. and $H\|$w.s., respectively. $R(H)$ in the mixed state for a given J is lower than for the wire specimen [see Fig. 1(a)]. Of the two field orientations, it is smaller when $H\|$w.s. The anomalous resistance minimum is also smaller for this field direction. While H_{c2} is independent of field direction, H_{c3} has a higher value when $H\|$w.s. In either case, H_{c3} tends to decrease as J increases. If H_{c3} is evaluated at the smallest J, namely $J=10$ A/cm², the ratio $H_{c3}/H_{c2}=1.68$, when H is parallel to the wide side. This value is in good agreement with the Saint-James and de Gennes analysis[21] which predicts $H_{c3}/H_{c2}=1.695$, when H is parallel to the surface. For H perpendicular to the surface, the Saint-James and de Gennes analysis reduces to that of Abrikosov with no remanent superconductivity expected beyond H_{c2}. In this study when the field is perpendicular to the wide side of the annealed ribbon [Fig. 7(a)] some remanent superconductivity does exist and $H_{c3}/H_{c2}=1.37$ ($J=10$ A/cm²). The ratio is smaller than the theoretical value, although the edges of the ribbon specimen provide some surface parallel to the field.

b. Nb+0.70 at.% Oxygen-Cold-Worked Ribbon. When a ribbon, the same geometry as that above, is made by cold-working niobium wire containing 0.70 at.% oxygen (i.e., wire uniformly compressed to a flat ribbon) there occurs a decrease in the resistance in the mixed state and an enhancement of the resistance

[21] W. DeSorbo, Rev. Mod. Phys. 36, 90 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

minimum for either of the two field directions when compared with the results obtained on the annealed specimen [compare Fig. 7(a) with 8(a) and 7(b) with 8(b)]. The resistance decrease is more pronounced when $H\|$w.s. of the ribbon. This decrease in $R(H)$ in the mixed state upon cold working is consistent with the point of view[16] that R is determined from the rate at which flux filaments are moving. Increasing dislocation (or defects) and/or segregation and precipitate concentration impedes the motion decreasing R. The Nb–O solutions, as prepared, are all supersaturated at room temperature (equilibrium solubility at room temperature is vanishingly small.)[13,14] Upon cold working, the dislocations can help take oxygen out of solution since they can serve as sites for segregation and precipitation and provide paths for enhanced diffusion.

The resistive anomaly and its enhancement by cold work may be related to the relative orientation between the flux filaments in the mixed state and defects with that of the field.[26] On the basis of the Anderson and Kim[20] suggestion, flux line rigidity would not only increase near H_{c2} with cold work, but it would also be anisotropic. With the type of deformation used in this work, this relative orientation would be expected to be a complex one. The demagnetization coefficient itself seems to be important in the low-field region. The small change, if any, of H_{c2} upon cold work (see Fig. 9) implies, from the Abrikosov model, that the mean free path of the bulk specimen has not been altered appreciably and that possibly only a small fraction of the total solute atom segregating to defects are responsible for the large changes in the resistive characteristics observed in the mixed state. The small changes in H_{c2} with cold work accompanied by large changes in magnetization peak and hysteresis loop are consistent with earlier observations made on some lead-base alloys.[27]

For cold-worked ribbon, H_{c3} is larger than that observed on the annealed ribbon, and the ratio H_{c3}/H_{c2} equals 1.91 ($J=111$ A/cm^2) for either field direction [see Figs. 8(a) and (b)] and is expected to be even larger for smaller J. The increase in H_{c3} is anticipated if defects participate in the nucleation to the superconducting sheath as suggested by Saint-James and de Gennes. The distinction in $R(H)$ between parallel and perpendicular field orientations for the cold worked ribbon is less apparent in the region H_{c2} to H_{c3} [compare Figs. 8(a) and (b)]. These observations are in agreement with similar results recently reported by Hempstead and Kim[23] on some substitutional solid solutions of Type II superconductors.

c. Dependence of H'—Nb+0.70 at.% Oxygen (An-

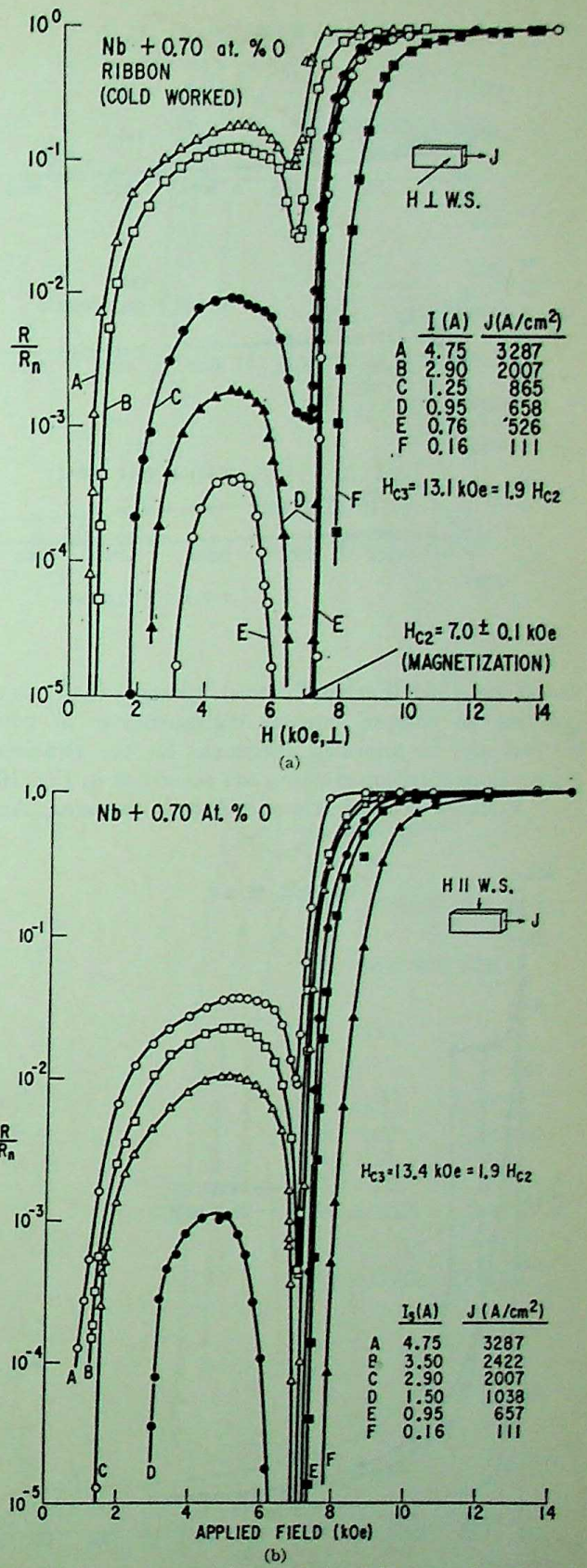


FIG. 8. The resistance characteristics R/R_n as a function of applied transverse field at 4.2°K for a cold-worked ribbon (0.032×0.007 in.) containing 0.70 at.% oxygen (a) with field perpendicular to wide side; (b) with field parallel to the wide side. H_{c3} values were evaluated at $J=111$ A/cm^2.

[26] Anisotropic effects observed in current density-field studies of some substitutional alloys [J. J. Hauser and R. G. Treuting, Phys. Chem. Solids 24, 371 (1963)] have been related to the presence of an "anisotropic filamentary structure" produced by cold rolling and not specifically to changes in the demagnetization coefficient.

[27] J. D. Livingston, Phys. Rev. 129, 1943 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

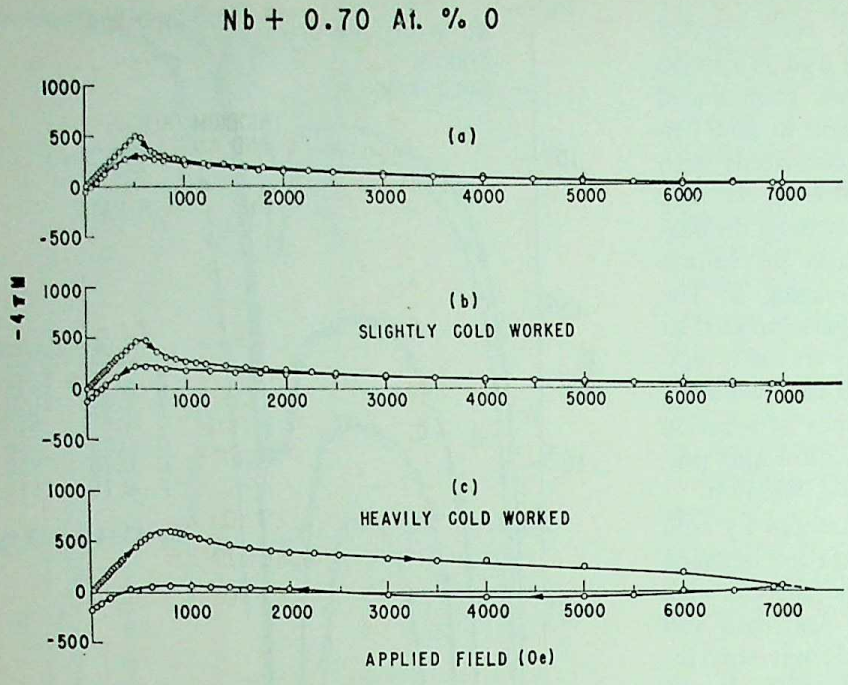


FIG. 9. The influence of cold work on the magnetization behavior at 4.2°K of niobium +0.70 at.% oxygen (a) 0.030-in. diam wire, (b) same wire with diameter reduced to ~0.027 in. by hammering, and (c) wire after uniformly compressed to flat ribbons 0.032×0.007 in.

nealed and Cold-Worked Ribbons). Changes in H' as a function of sample current corresponding to cold worked and to annealed specimens for the two-field orientations mentioned above are presented in Fig. 10.

d. Niobium Ribbon—Outgassed and Annealed. Annealed niobium ribbon, whose interstitial concentration is less than 5 ppm, does not exhibit the resistance anomaly when the field orientation is either perpendicular or parallel to the wide side [Figs. 11(a) and (b)]. The resistance in the mixed state below H_{c2} (see also Fig. 12 for H_{c2} evaluated from the magnetization data) is larger when the field is perpendicular to the wide side.

Above H_{c2}, the resistance is smaller when $H\|$w.s. of the ribbon. This could be due to the existence of a superconducting sheath predicted for this field orienta-

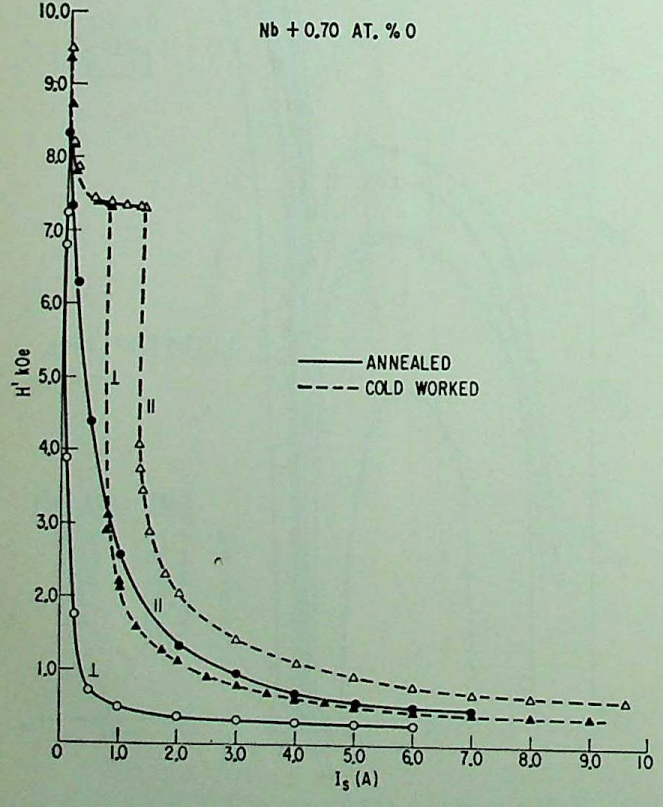


FIG. 10. The transverse field H' required for the initial appearance of resistance ($R/R_n<1\times10^{-3}$) at 4.2°K as a function of specimen current, I_s (amperes) for (a) annealed and (b) cold-worked ribbon of Figs. 7 and 8.

TABLE I. The influence of oxygen concentration in niobium wires (0.030-in. diam) on the ratio H_{c3}/H_{c2} at 4.2°K (see note added in proof). Magnetic field perpendicular to wire and to flow of current.

	T_c(°K) (Ref. 2)	H_{c2}[a] (kOe)	K[b]	H_{c3}/H_{c2} ($J=10$ A/cm²)
Nb				
wire	9.4_6	2.70	1.24	2.5_2
ribbon $H\|$w.s.	9.4_6	2.70	1.24	2.4_4
ribbon $H\perp$w.s.	9.4_6	2.70	1.24	1.9_0
Nb+200 ppm O wire				2.0_0
Nb+0.7 at.%O wire	8.7_8	7.0	3.64	1.6_0
ribbon $H\|$w.s.	8.7_8	7.0		1.6_8
ribbon $H\perp$w.s.	8.7_8	7.0		1.3_7
Nb+1.52 at.%O wire	8.0_4	9.6	6.03	1.2_1
Nb+6.4 at.%O wire	~9.0			<1.2

[a] Determined magnetically.
[b] For value of H_c used to evaluate K, see Ref. 2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

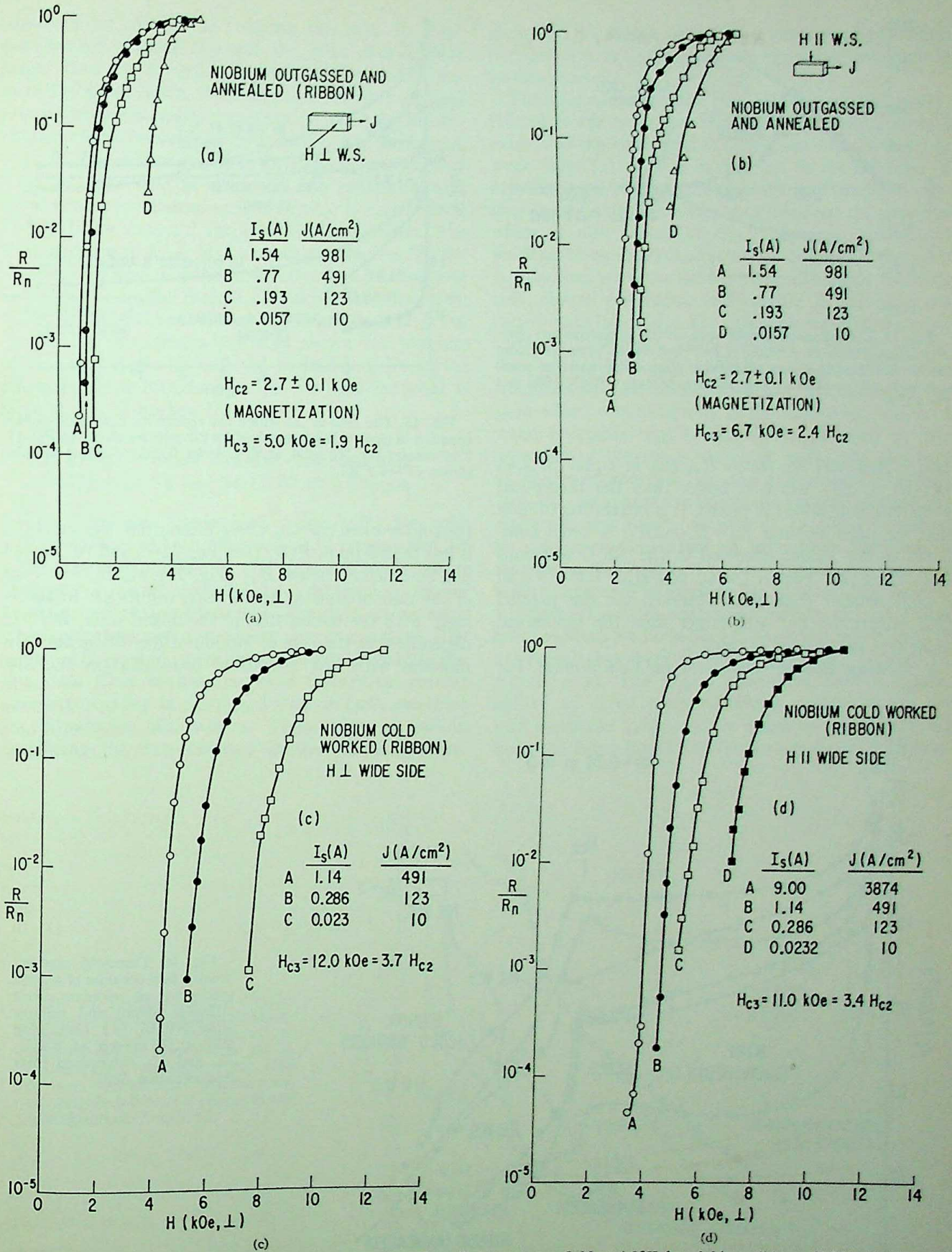


FIG. 11. The resistive characteristics R/R_n as a function of applied transverse field at 4.2°K for niobium ribbons (0.032×0.007 in.) containing less than 5 ppm interstitials: (i) annealed; (a) field perpendicular to wide side (b) parallel to wide side; (ii) cold worked by uniformly compressing a 0.030-in. diam wire (ribbon diamensions 0.032×0.007 in.); (c) field perpendicular to wide side; and (d) field parallel to wide side. H_{c3} values are evaluated at $J \cong 10$ A/cm².

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

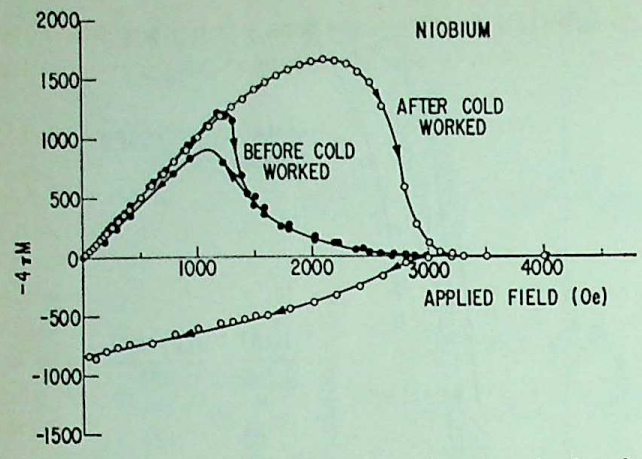


FIG. 12. The influence of cold work on the magnetization characteristics of niobium containing less than 5 ppm oxygen content for an annealed specimen (0.030-in. diam wire) and the same specimen cold worked by compression (ribbon 0.032×0.007 in.) $T=4.2°K$.

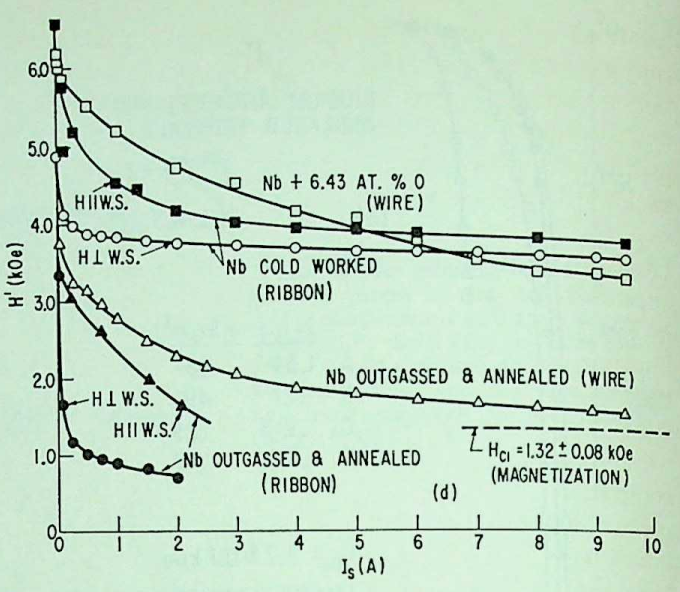


FIG. 13. The field H' at which the resistivity first appears as a function of current is summarized for Nb ribbons shown in Fig. 11. The results for Nb+6.4 at.% O (wire 0.032-in. diam) are also shown ($T=4.2°K$).

tion.[21] However, the ratio H_{c3}/H_{c2} is equal to 2.44 ($J=10$ A/cm²) which is larger than the theoretical value (see note added in proof). It is interesting to note that for pure niobium $K(=H_{c2}/\sqrt{2}H_c)$ is small compared to the values observed in the Nb–O solutions (see Table I). For $H\perp$w.s., $H_{c3}/H_{c2}=1.85$ ($J=10$ A/cm²) smaller than that observed for the parallel field orientation but still larger than the theoretical value (see note added in proof).

e. Niobium Ribbon—Outgassed and Cold Worked. For the cold-worked ribbon, whose interstitial concentration is less than 5 ppm, $R(H)$ [see Fig. 11(c) and (d)] could not be detected below H_{c2} (Fig. 11) within the limits of the measurements. Again, such results are in agreement with the notion that in the mixed state, $R(H)$ is dependent on the rate of motion of flux filaments.[16] In niobium with low interstitial concentration the dis-

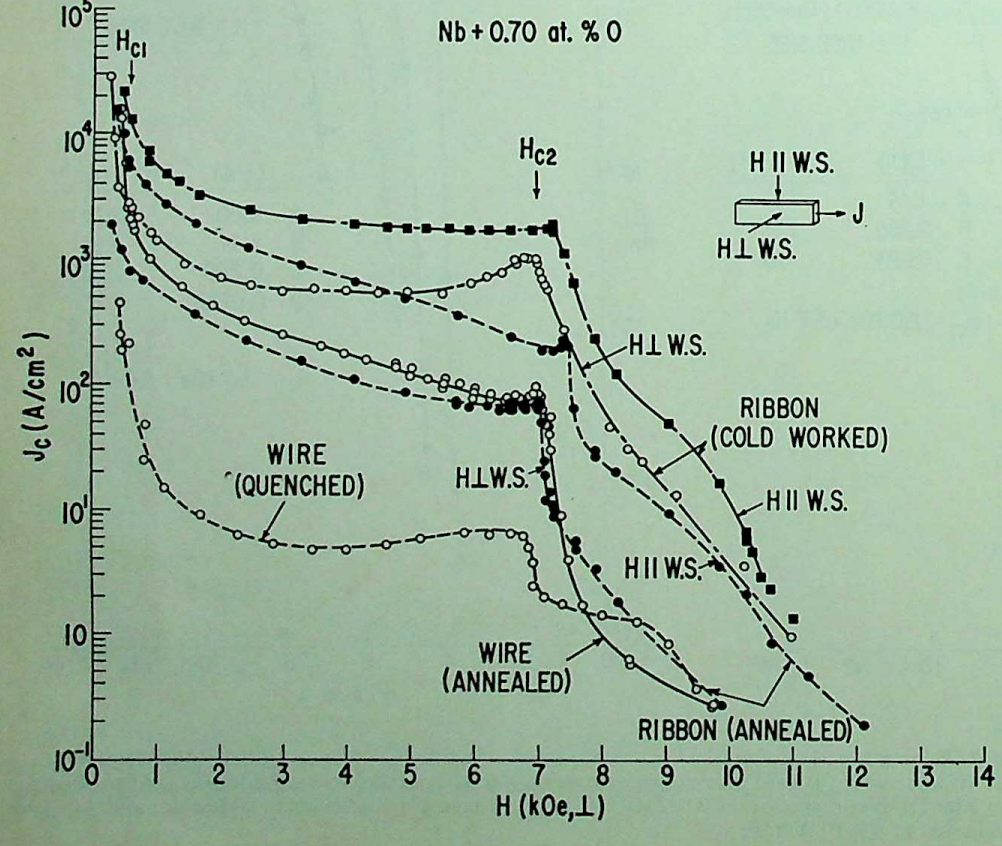


FIG. 14. Threshold current-density field behavior of several specimens of niobium wire (0.030-in. diam) and ribbons (0.032×0.007 in.) containing 0.70 at.% oxygen at 4.2°K. H_{c1} and H_{c2} determined from magnetization data.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

locations are effective in reducing this rate. In these cold-worked ribbons H_{c3} and the ratio H_{c3}/H_{c2} are larger than the corresponding values observed in annealed ribbons. Again, the distinction between the field orientation of the resistance transition is less significant, similar to the behavior in cold-worked Nb–O ribbons.

f. H' Dependence on Field Orientation in Niobium (Ribbon). The field H' at which the resistivity first appears as a function of sample current is summarized in Fig. 13 for annealed and cold-worked niobium. The limiting value of H' at high current approaches H_{c1} for the annealed wire. For the annealed ribbon, the limiting value may be smaller than H_{c1} for either field direction, while for cold-worked ribbon it is larger than H_{c1}. For comparison, the results are also shown for niobium containing oxygen beyond the solubility limit at the temperature of introduction. H' for this material is large as in cold-worked Nb.

2. Critical Current Density—Annealed and Cold-Worked Nb and Nb+0.70 at.% Oxygen

The threshold current-field data $J_c(H)$ for cold-worked ribbons of both Nb+0.70 at.% oxygen and Nb are summarized in Figs. 14 and 15, respectively, where they are compared with the data obtained on similar but annealed ribbon specimens. The data on annealed wire specimens are also shown for comparison. The values of H_{c1} and H_{c2} were obtained from the magnetization data. Cold working increases the critical current carrying capacity in the mixed state in both materials by introducing dislocations. These play a role in stabilizing the flux filaments in the mixed state reducing $R(H)$. In Nb-O solutions, some segregation and precipitation at these dislocations may be participating in this process.

The anomolous increase in $J_c(H)$ curve (peak effect) related to the resistance minimum, discussed above, is evident in the interstitial solid solution even before cold work (Fig. 14). This is in contrast to the behavior in substitutional solid solutions where it is usually necessary to cold work the material to bring out the anomaly where it also occurs at H_{c2}.[25] However, upon cold working these interstitial solutions (ribbons), the peak is enhanced when the field is perpendicular to the wide side, always coincident with smaller J_c values in the mixed state below H_{c2}. For this same material the peak virtually disappears when the field is parallel to the flat surface, but $J_c(H)$ is now larger in the mixed state. Mendelssohn (see discussion of Ref. 25) attributes the peak effect, occurring only in the transverse field, to a more favorable relative orientation assumed by the direction of the superconducting paths and the field as H_{c2} is approached (see also Hauser and Treuting, Ref. 26). An alternative explanation is that proposed by Anderson and Kim[20] (see above). In either outgassed or cold-worked niobium no anomalous variation is evident in the $J_c(H)$ curve for either of the two field orientations used to examine the ribbons (Fig. 15).

Above H_{c2} field orientation effects on $J_c(H)$ are quite evident for both Nb+0.70 at.% oxygen and Nb after each has been annealed (see Fig. 14 and 15, respectively). The ribbon specimens show a greater ability to carry superconducting current when the field is parallel to the surface which would be expected from the Saint-James and de Gennes model.[21] In cold-

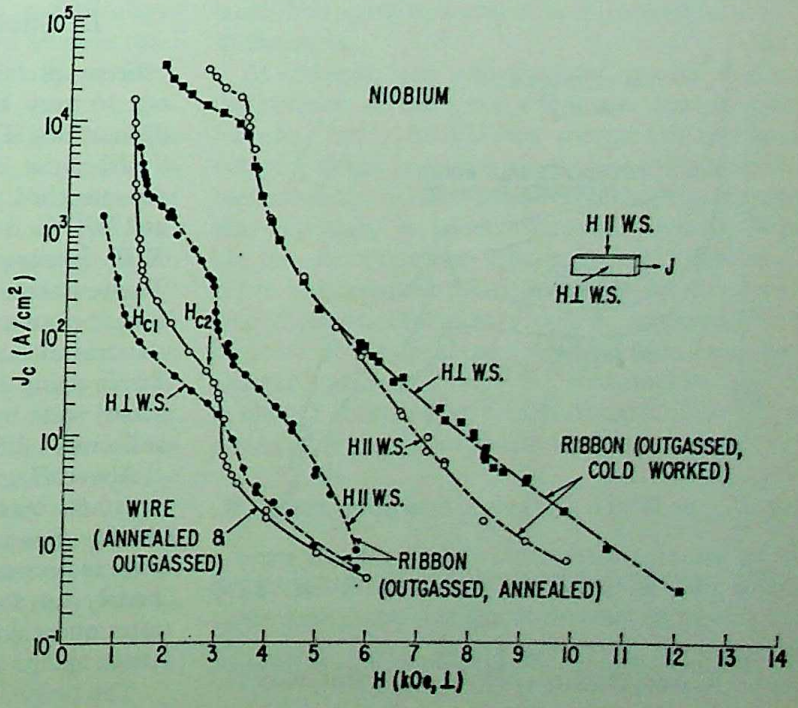


FIG. 15. Threshold current density-field behavior of several wire (0.030-in. diam) and ribbon (compressed to 0.030×0.007 in.) specimens of niobium, each containing less than 5 ppm interstitial concentration, showing the influence of cold work and field direction (T=4.2°K).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

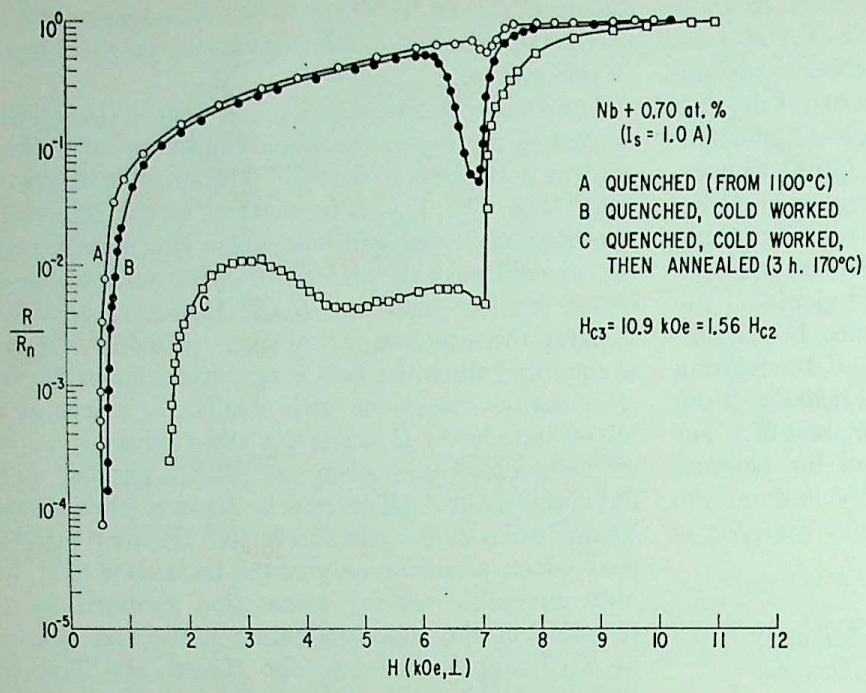


FIG. 16. Resistance characteristics R/R_n of a wire specimen (0.030-in. diam) of Nb+0.70 at.% oxygen; (a) quenched in helium exchange gas from an homogenization temperature 1100°C to room temperature; (b) quenched and cold worked by uniform compression roughly to a rectangular form approximately 0.028×0.032 in.; (c) after (b) the same wire was given a low temperature anneal, 3 h at 170°C ($J=220$ A/cm^2) ($T=4.2$°K). At this current density $H_{c3}=10.9=1.56$ H_{c2}.

worked ribbons the field orientation effect on $J_c(H)$ is smaller and for Nb ribbon it is nonexistent just below H_{c2} with the deviations going over into the opposite direction for small $J_c(H)$ values. Similar results have been observed in cold-worked Nb+3 at.% Ti.[25] These results corroborate the dependence of resistance on field orientation above H_{c2} described above. As the defect concentration increases, the distinction between $R(H)$ for $\parallel$ and $\perp$ field orientation becomes less significant.

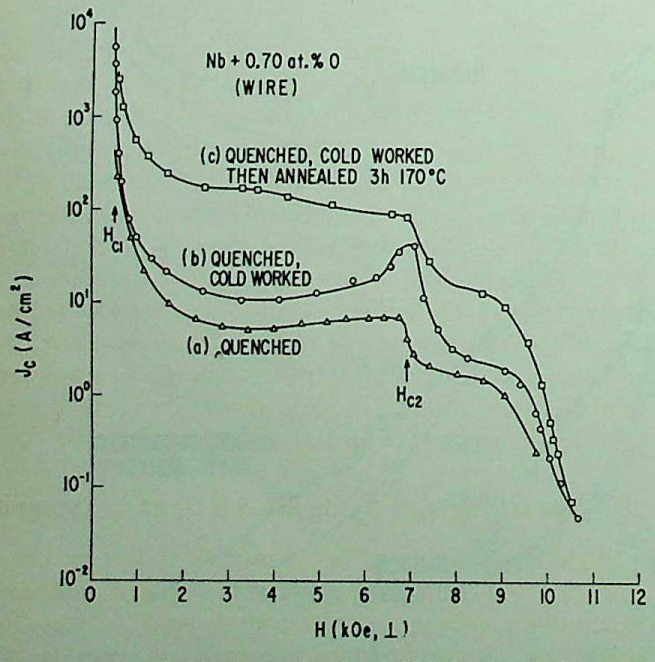


FIG. 17. Threshold current-density field behavior of the same specimens shown in Fig. 16 ($T=4.2$°K).

C. Ratio H_{c3}/H_{c2} and Oxygen Concentration in Niobium

The ratio H_{c3}/H_{c2} observed on annealed niobium wires containing various concentrations of oxygen are summarized in Table I. The values are approximately equal to those observed on ribbon samples when $H\parallel$w.s. (e.g., Nb and Nb+0.70 at.% oxygen). The ratio is largest for pure niobium which also has the lowest K. It decreases with increasing K and increasing solute concentration (see note added in proof).

D. Strain Aging Nb+0.70 at.% Oxygen

Some preliminary experiments have been carried out to vary the degree of interstitial segregation to dislocations. This was done by studying both $R(H)$ and $J_c(H)$ on a wire specimen, Nb+0.70 at.% oxygen, (a) quenched, (b) then cold worked, and (c) annealed at 170°C for 3 h in a pressure less than 1×10^{-7} mm Hg. $R(H)$ data are summarized in Fig. 16 for $J=220$ A/cm^2. The resistance in the mixed state is approximately the same before and after cold working. However, the resistance anomaly in the latter is much more pronounced. Strain aging at 170°C decreases the resistance in the mixed state by a significant amount, masking out the resistance minimum.

Above H_{c2}, the data show little, if any, change in resistance upon quenching and cold work for the current density shown. The strain aging process does, however, tend to decrease R. H_{c3}/H_{c2} for the J shown is approximately 1.4_2 for these wire specimens, both before and after quenching; increasing to approximately 1.6 after strain aging.

The process of homogenizing, cold working and low

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

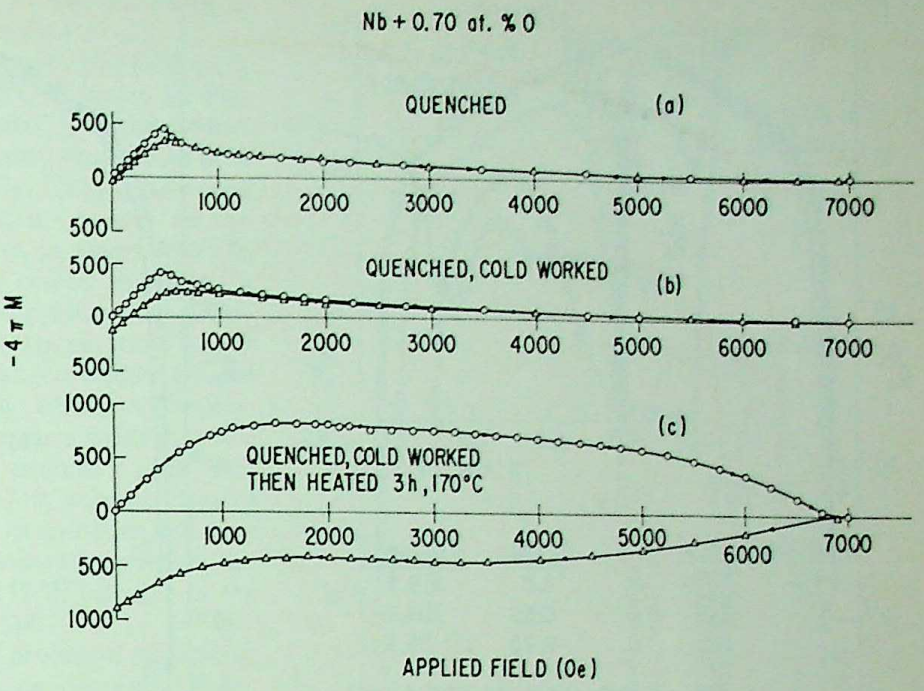


FIG. 18. The magnetization characteristics at 4.2°K of niobium containing 0.70 at.% oxygen; (a) 0.030-in.-diam wire after quenching in helium exchange gas from an homogenization temperature of approximately 1100°C; (b) a similar wire cold worked by uniformly compressing and reducing the wire to roughly a rectangle 0.040×0.019 in.; and (c) specimen (b) heated 3 h at 170°C.

temperature anneal on $J_c(H)$ and on the magnetization loop are summarized in Figs. 17 and 18, respectively. It is interesting to note that strain aging brings $J_c(H)$ in the mixed state up to where it was for slow-cooled specimens (compare Fig. 17 with Fig. 14). This result cannot be deduced from the corresponding magnetization loops (compare Fig. 16 with Fig. 8). It seems that structural changes brought about by cold working and by strain aging responsible for changes in $J_c(H)$ in the mixed state also cause changes above H_{c2} (see Fig. 17). The results indicate that in these specimens, if a superconducting sheath is carrying the current above H_{c2}, this sheath appears to be dependent on structural changes. Below H_{c1}, the current, carried in the penetration depth, appears to be independent of these changes.

The results, both $R(H)$ and $J_c(H)$, suggest that the anomalous peak effect, or resistance minimum, is strongest at some intermediate stage of the process of segregation and precipitation at dislocations. This conclusion is in agreement with the suggestion made earlier[25] that a solute or impurity atom is probably necessary for the appearance of the anomaly in both interstitial as well as substitutional solid solutions. The more wide-spread appearance of the anomaly in the former solutions has been attributed[25] to the far greater mobility (e.g., see Ref. 28) of the interstitial atom. This makes the interstitial solute atom segregate more easily to structural irregularities, such as dislocations.

As mentioned above, the anomaly is readily evident in niobium containing as little as 200 ppm oxygen, but not in Nb containing less then 5 ppm oxygen. Recent transmission electron microscopy studies[3] have shown that the presence of relatively small amounts of interstitial impurity (probably as little as 150 ppm) has a strong influence on the annealing structures of the metal. Dislocations were reported to act as sinks for interstitials present as interstitial atmospheres or precipitates. These would also tend to retard the rate of the recrystallization process and probably account for a wide variety of structures seen after high-temperature anneals. The appearance of the anomaly in niobium reported earlier[4-6] was probably due to the presence of small but significant amounts of interstitial impurities in the metal.

At present, the experimental results outline the dependence of the peak effect on various structural features, but a satisfactory theory has not been developed. Some proposed alternative explanations have been cited above. Hauser and Treuting[26] have suggested that anisotropy in defect structure is probably responsible for the minimum ("valley effect") in the $J_c(H)$ curve but reported no evidence for the presence of a second phase. Preliminary optical microscopy studies on some of the specimens reported here revealed precipitate particles only in cold-worked specimens (ribbons) occurring at a high density in certain areas presumably regions of greater dislocation density.

E. Quenching and Aging Nb+0.70 at.% Oxygen

Since strain aging at temperatures as low as 170°C has a marked effect on the mixed state, some additional experiments were carried out to study aging of quenched specimens. Some preliminary work on varying the homogenization or solution temperature, T_h was also

[28] G. B. Gibbs, D. Graham, and D. H. Tomlin, Phil. Mag. 8, 1269 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

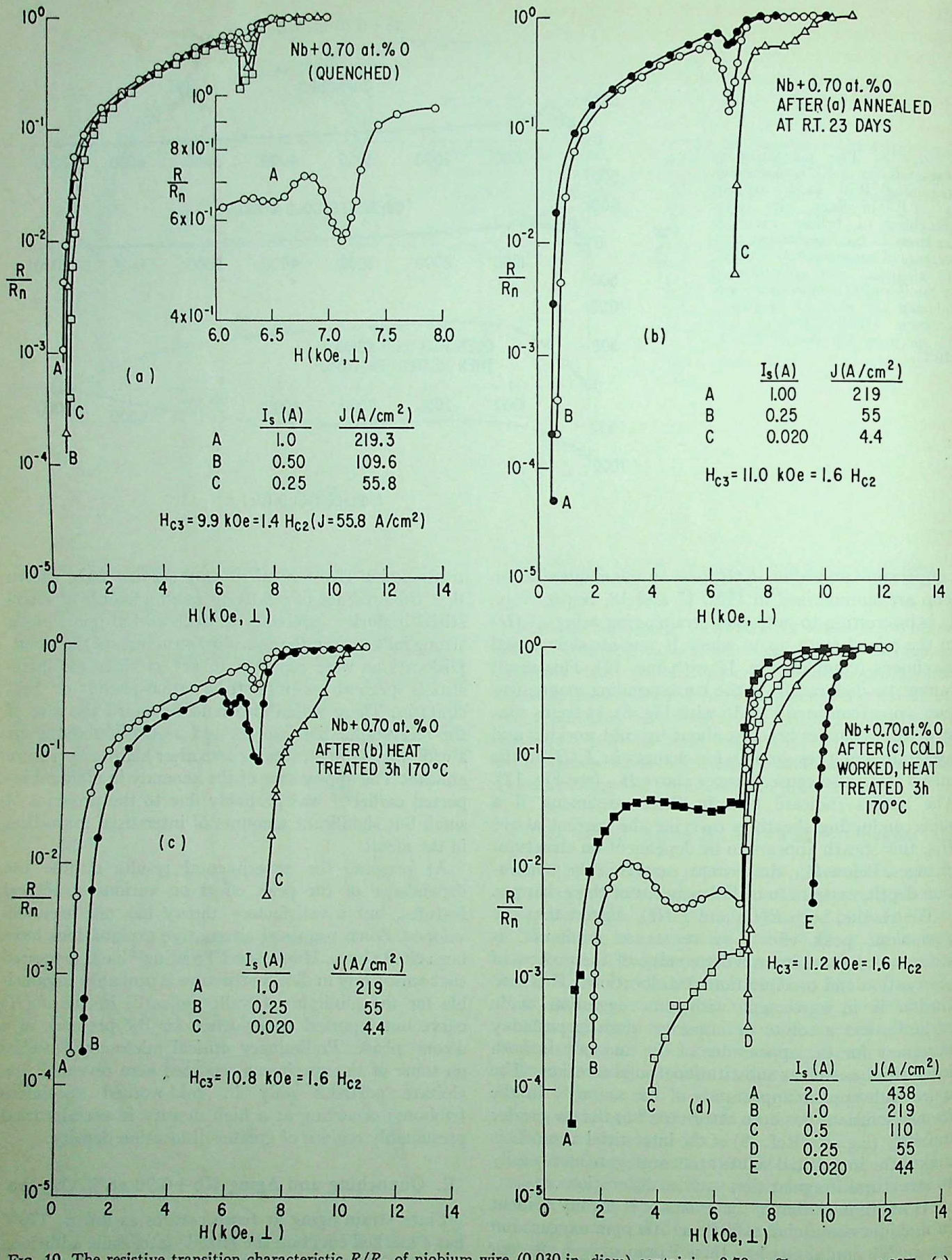


FIG. 19. The resistive transition characteristic R/R_n of niobium wire (0.030-in. diam) containing 0.70 at.% oxygen at 4.2°K; (a) after quenching in helium exchange gas from an homogenization temperature of approximately 1000°C; H_{c3} evaluated at J=55.8 A/cm^2; (b) annealed an additional 23 days at room temperature; (c) then heated 3 h at 170°C; (d) after (c) cold work by uniform compression reducing the diameter from 0.030 in. to ~0.027 in. then heated 3 h at 170°C. In (b), (c), and (d) H_{c3} evaluated at J=4.4 A/cm^2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

attempted. In these experiments T_h was decreased to liquid nitrogen temperature in less than 20 sec.

The resistive transition of a sample Nb+0.70 at.% oxygen quenched from $T_h=1000°C$ is shown in Fig. 19(a). When compared to a similar specimen cooled from the same temperature to room temperature at a slower rate (approximately 1 h) [Fig. 1(a)], it is evident that increasing the supersaturation, as well as the degree of disperseness of the solute, increases $R(H)$ in the mixed state. In the rapidly cooled specimen, a secondary dip is evident below the primary one occurring at H_{c2}. Upon aging at room temperature for 23 days, $R(H)$ remains essentially unaffected, although the secondary dip has disappeared [Fig. 19(b)]. A low temperature anneal, 3 h at 170°C, reveals a small decrease in $R(H)$, a reappearance of the secondary peak, and definite changes above H_{c2} at low current density [Fig. 19(c)]. Upon cold working the specimen and then annealing (3 h at 170°C), pronounced decreases are evident in the mixed state [Fig. 19(d)] similar to the strain aging results shown in Fig. 16.

In the region $H_{c2}\lesssim H\lesssim H_{c3}$ the process of quenching, aging, annealing, and strain aging decreases $R(H)$. Structural changes responsible for resistance changes in the mixed state are also responsible for changes observed above H_{c2}. The results again emphasize the importance of segregation and precipitation at dislocations on reducing resistance both in the mixed-state region as well as in the region above H_{c2}. If the superconducting sheath proposed for Type II superconductors[21] is responsible for the variations of $R(H)$ above H_{c2}, then the sheath is influenced by the segregation and precipitation at dislocations. However, for these wire specimens, H_{c3} ($J=4.4$ A/cm^2) appears to be unaffected by the various structural changes introduced. The ratio H_{c3}/H_{c2} remains constant with a value approximately equal to 1.6.

Upon increasing the homogenization temperature from 1000 to 1200°C and quenching rapidly, $R(H)$ in the mixed state tends to remain unchanged but the resistance minimum is now larger [compare Fig. 20(a) with Fig. 19(a)]. Increasing the homogenization temperature does not appear to change the ratio, $H_{c3}/H_{c2}=1.6$. However, in this case the resistive transitions at low current density reveal some distinct secondary minima between H_{c2} and H_{c3}. These secondary anomalies may be a consequence of some strains introduced in the wire upon cooling from higher T_h which provide new diffusion paths resulting in some regions having a slightly higher solute concentration with a corresponding higher H_{c2}. Another possibility is that these new variations are due to differences in structure that may exist at the surface and the interior of the niobium matrix when the metal is deformed a relatively small amount.[3] This different surface structure may also have a higher H_{c2}. Upon heating this specimen at 170°C for 3 h, the minima above H_{c2} disappear [Fig. 20(b)]. A relatively larger decrease in

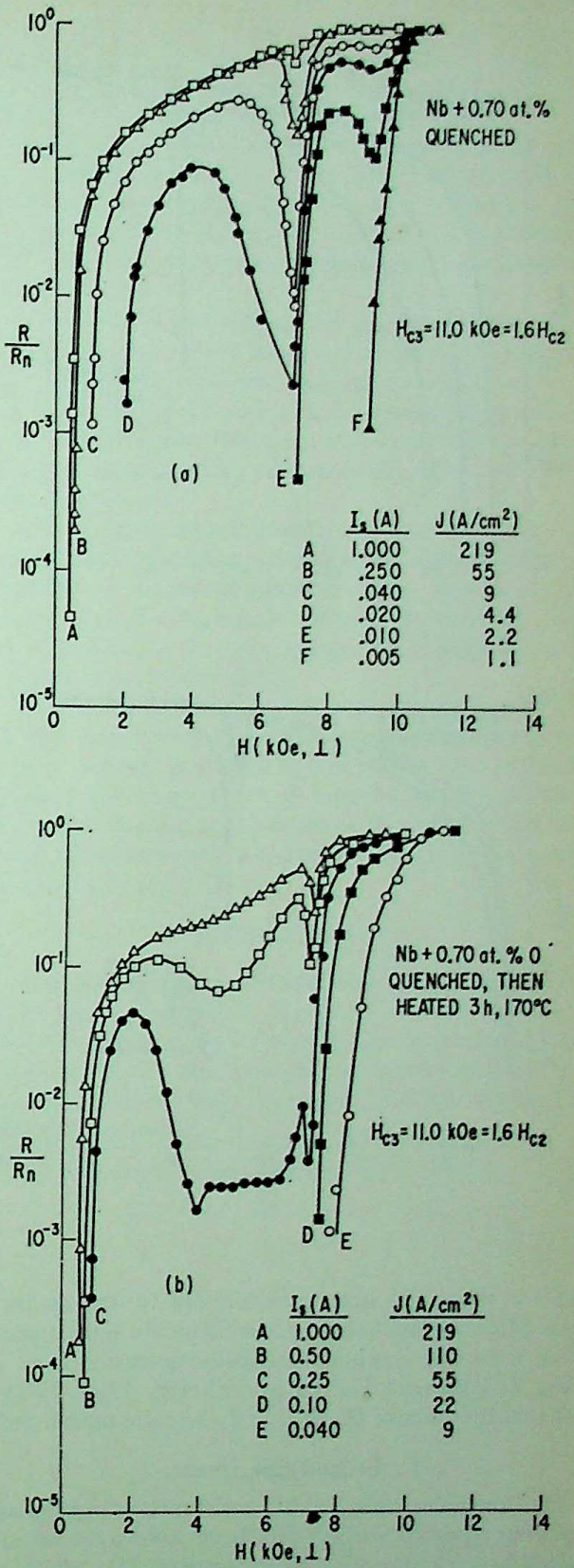
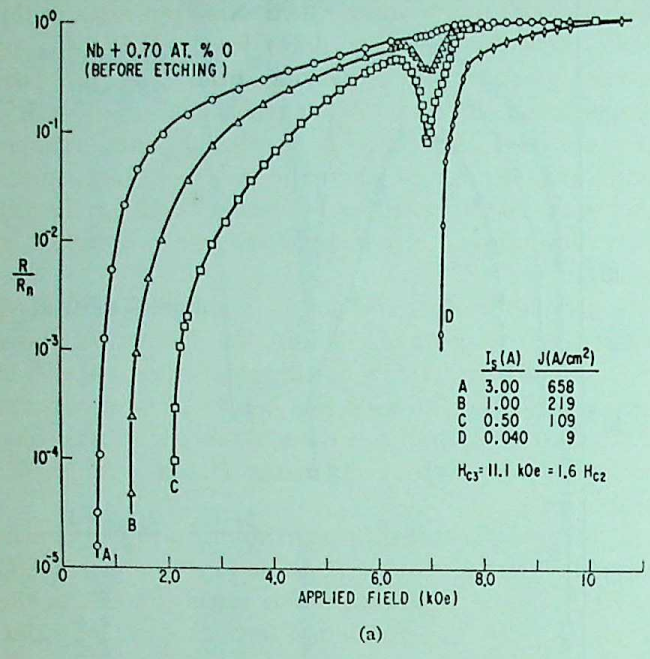


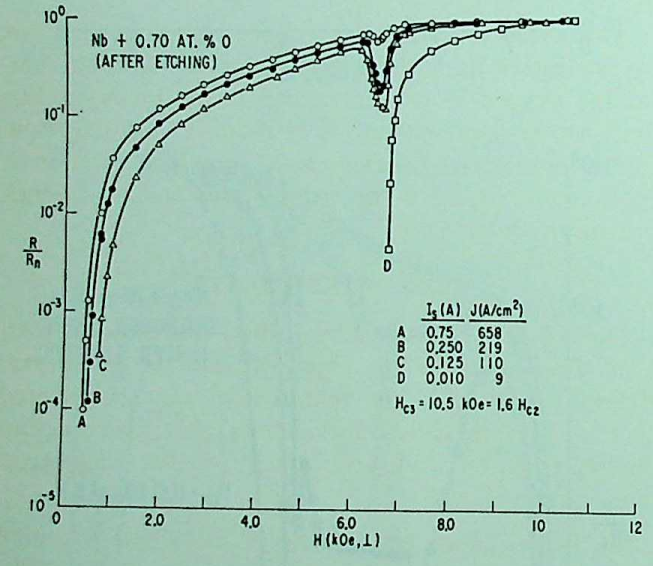
FIG. 20. The resistance characteristics R/R_n of niobium wire (0.030-in. diam) containing 0.70 at.% oxygen ($T=4.2°$K); (a) quenched in helium exchange gas from an homogenization temperature of approximately 1200°C; (b) then heated 3 h at 170°C. H_{c3} evaluated at $J=10$ A/cm^2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

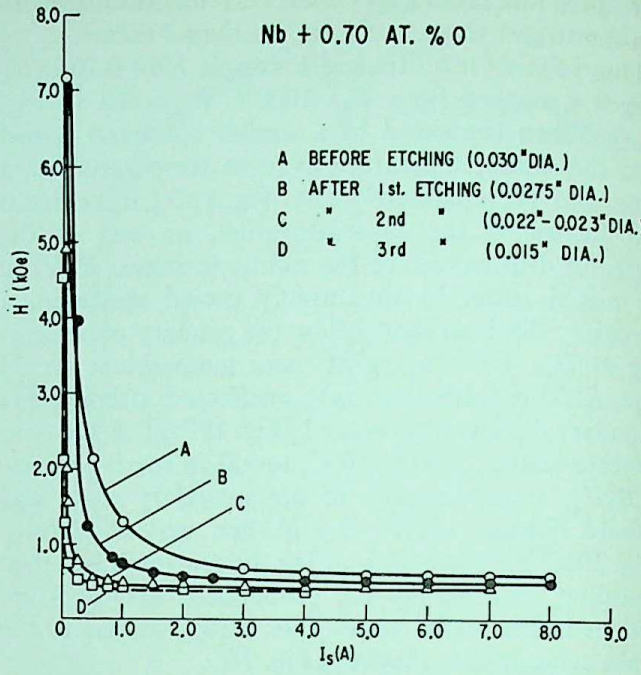
Digitized by Arya Samaj Foundation Chennai and eGangotri



(a)



(b)



(c)

FIG. 21. The resistive characteristics R/R_n as a function of applied transverse field at 4.2°K; (a) before etching specimen of niobium +0.70 at.% O, wire (0.030-in. diam); (b) after etching diameter reduced from 0.030 to 0.015 in. diam; (c) the field H' required for the initial appearance of resistance ($R/R_n \lesssim 1\times10^{-3}$) at 4.2°K as a function of specimen current I_s (amperes) before and after etching.

$R(H)$ in the mixed state, probably due to strain aging, takes place over that observed for an identical treatment given a similar specimen originally quenched from a lower T_h [compare Fig. 20(b) with Fig. 19(c)]. Note also that the value of H_{c3} or H_{c3}/H_{c2} remains unchanged.

F. Etched Specimens

The anomalous dip in the resistance transition of niobium reported earlier[6] has been known to be influenced by etching (see discussion Ref. 25). Nb–O or Nb–N specimens prepared in this work when etched with a solution consisting of 8 HNO_3, 2 H_2O, 1 HF showed only little changes in the dip characteristics even when the diameter was reduced by one half [compare Figs. 21(a) and (b)]. For example for $J=110$ A/cm², the field at which the minimum occurs, changes from 6.95 kOe before etching to 6.75 kOe after etching. This small decrease indicates only a small change in the average solute concentration. Assuming that changes in H_{c2} reflect average changes in solute concentration, the experiment could be useful in investigating diffusion of nonmetallic solutes in superconducting transition metals. After etching a significant increase in $R(H)$ in the mixed state, however, does occur, indicating less effective pinning of flux filaments. Above H_{c2}, the influence of etching on $R(H)$ is small or negligible. H_{c3}/H_{c2} remains unchanged and equal to approximately 1.6.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

H' as a function of specimen current exhibits a decreasing trend upon reducing the wire diameter by etching [Fig. 21(b)]. Its limiting value may be smaller than H_{c1}.

SUMMARY

1. Oxygen and nitrogen as interstitial solutes in niobium cause the metal in transverse field to exhibit anomalous resistance minima or critical current maxima (peak effect). The anomaly is evident in niobium containing as little as 200 ppm oxygen but not in niobium whose interstitial concentration is approximately 5 ppm.

2. The anomaly appears at H_{c2}, the upper critical field of these Type II superconductors determined magnetically.

3. The anomaly is increased by cold working but decreased by strain aging (3 h, 170°C) and by rapid quenching. These results suggest that the anomaly is strongest at some intermediate state of the process of segregation and precipitation at dislocations.

4. The resistance (R) in the mixed state is decreased by exceeding the solute solubility limit, by cold working and by strain aging at 170°C; it is increased in the mixed state by introducing more interstitial solute, by quenching and by etching.

5. Small changes, if any, of H_{c2} upon cold work seem to indicate that only a small fraction of the total solute atom segregating to defects are responsible for large changes in both the resistive and critical current characteristics.

6. For an annealed ribbon of niobium containing 0.70 at.% oxygen ($K=3.6$) and for field orientation parallel to the wide side, H_{c3}/H_{c2} gradually increases for decreasing current density reaching a value 1.68 for $J=10$ A/cm^2, in good agreement with the predicted value of Saint-James and de Gennes. For transverse field, H_{c3}/H_{c2} is equal to 1.37.

7. For an annealed niobium ribbon containing less than 5 ppm oxygen ($K=1.2_4$) for field orientation parallel to the wide side, $H_{c3}/H_{c2}=2.44$ ($J=10$ A/cm^2), larger than the predicted value. The value is approximately equal to the ratio observed for an *annealed* niobium *wire* in a *transverse* field. Niobium wires in this field orientation show a dependence of H_{c3}/H_{c2} on oxygen concentration, varying from 2.44 for pure niobium ($K=1.2_4$) to 1.21 for niobium containing 1.52 at.% oxygen ($K=6.0_3$). It is possible that inhomogeneities in the sample where the surface is characterized by one K and the interior by another could artifically force this ratio in either direction (see note added in proof).

8. In cold-worked ribbons of niobium +0.70 at.% oxygen, or cold-worked niobium, the distinction, above H_{c2}, in $R(H)$ or $J_c(H)$ for field orientations parallel or perpendicular to the wide side becomes less significant than the corresponding values observed for annealed ribbons. Cold working increases H_{c3}/H_{c2} for either field orientation.

9. Structural changes brought about by rapid quenching, aging, strain aging responsible for changes in $J_c(H)$ or $R(H)$ in the mixed state also cause changes in the region $H_{c2}<H<H_{c3}$ but do not seem to affect the value of $H_{c3}=1.6\ H_{c2}$ for wire specimens (Nb+0.70 at.% oxygen) in transverse fields.

Note added in proof. Subsequent results carried out by the author on both pure niobium and niobium containing oxygen in concentrations below the solubility limit at the temperature of introducing oxygen show the value H_{c3}/H_{c2} is independent of κ. The value is in good agreement with the Saint-James and de Gennes theoretical value 1.695.[29]

ACKNOWLEDGMENTS

The author has benefitted from discussions with J. D. Livingston, H. R. Hart, Jr., and P. S. Swartz. Helpful criticism of the manuscript was also offered by them. W. H. Healy provided capable assistance in the experimental work and carried out the magnetization measurements.

[29] W. DeSorbo, Bull. Am. Phys. Soc. 9, 253 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Scattering of Energy-Time Wave Packets from Many-Body Systems*

L. R. DODD
Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia

AND

I. E. MCCARTHY
Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia and Department of Physics, University of California, Davis, California

(Received 7 January 1964)

The wave-packet nature of an experimental beam is discussed. The theory of wave-packet scattering is applied to wave packets with exponential time dependence scattered from resonances in both the isolated-resonance and overlapping-resonance regions. It is shown how a sequential description of scattering depends on some of the resonance parameters in the scattering amplitude. The meaning and usefulness of some experimentally possible wave-packet experiments is discussed.

1. INTRODUCTION

SCATTERING of a particle from a many-body system can be divided into two types with respect to a time interval τ and a corresponding energy interval $\delta(\sim\hbar/\tau)$.[1] If the scattering amplitude as a function of energy varies rapidly over the energy interval δ, the scattering is said to be compound scattering; if the amplitude varies slowly over this interval, we have direct scattering. τ is several orders of magnitude greater than the time it takes the incident particle to traverse a distance the size of the scatterer in free space.

In most actual scattering experiments time is not resolved. The incident beam has an energy spread. The cross section can be regarded as the weighted sum of cross sections for independent experiments, each with definite incident energy. This will be discussed in Sec. 2. It is still possible to define a time delay Δt according to the definition of Wigner,[2]

$$\Delta t = -i\hbar(d/dE)\ln S(E), \tag{1}$$

where $S(E)$ is the S-matrix element for the scattering problem. This time delay may be thought of as the time it takes for the phase of one incident wave to catch up with that of another whose energy differs from it by an infinitesimal amount, and whose phase shift in the scatterer is therefore infinitesimally different.

If the S-matrix element is divided according to some physical prescription into a rapidly and a slowly varying component, this definition is difficult to apply and it certainly does not tell us anything about each component separately. The rapidly varying component corresponds to a large Δt, that is, to particles that spend a long time on the average in the scatterer, and the slowly varying component corresponds to a small Δt, that is to particles that pass rapidly over the scatterer without significant time delay.

An experiment which measures the relative amount of direct and compound scattering has been suggested by Eisberg, Yennie, and Wilkinson.[3] The experiment also defines an energy interval which is the reciprocal of the time delay. The information is obtained from the energy spectrum of bremsstrahlung from elastic scattering of charged particles.

One would like to make a classical picture in which particles are described as being in the scatterer for varying times. To do this one must make a wave-packet argument. It will be shown in Sec. 2 that the physical situation corresponding to a meaningful wave-packet argument is necessarily a time-dependent scattering experiment.

Wave-packet arguments are made for example by Friedman and Weisskopf[4] for the case of shape-elastic and compound-elastic scattering. They show that for nonoverlapping levels (resonances) of the compound system the time delay is the average over the beam energy spread of the time delays for the individual levels. If time delay is defined according to Wigner's definition, this result is obtained at once. It has been done explicitly for example by Goldberger and Watson.[5]

If, however, time delay is defined as the average time delay in the emerging of a wave packet in a hypothetical time-dependent experiment complementary to the usual energy-dependent experiment, then the result is not so clear. The delayed wave packet would be expected to interfere with the immediately scattered wave packet from the shape elastic scattering. It is commonly stated that if the two wave packets are sufficiently short in time they will not interfere.[6] This

* Supported in part by the U. S. Atomic Energy Commission and the Australian Commonwealth Scientific and Industrial Research Organization.

[1] For example H. Feshbach, C. E. Porter, and V. F. Weisskopf, Phys. Rev. 96, 448 (1954).

[2] E. P. Wigner, Phys. Rev. 98, 145 (1955).

[3] R. M. Eisberg, D. R. Yennie, and D. H. Wilkinson, Nucl. Phys. 18, 338 (1960).

[4] F. L. Friedman and V. F. Weisskopf, *Niels Bohr and the Development of Physics* (Pergamon Press, Ltd., London, 1955).

[5] M. L. Goldberger and K. M. Watson, Phys. Rev. 127, 2284 (1962).

[6] For example, H. G. Preston, *Physics of the Nucleus* (Addison Wesley Publishing Co., Inc., Reading, Massachusetts, 1962). R. K. Adair, S. E. Darden, and R. E. Fields, Phys. Rev. 96, 503 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

situation is now extreme hypothetical. It requires very good time definition in the experiment. In an actual experimental situation, time is undefined and the wave packets interfere completely. It may also be asked whether the scattering amplitude, and hence, for example, the angular distribution will not be dependent on the width of the incident wave packet, since, for short wave packets (in time), we might not expect such complete interference as for long wave packets.[7]

The question of what actually happens to a scattered wave packet has been investigated by Sasakawa.[8] He shows that for the scattering of a wave packet from an isolated resonance, the variation of the cross section with energy is characterized not by the width of the resonance but, in addition, by the width of the wave packet. In Sec. 3 we will re-derive this result using a different mathematical technique and Lorentzian wave packets. The Lorentzian shape is one shape for the energy spectrum that can be realized experimentally.

In Sec. 4 we will consider the scattering of wave packets from a system in which the compound levels overlap. This gives new insight into the optical model, which is defined for this purpose as the model for which the time delay in the complementary time-dependent experiment would be small. Our treatment is more general than Sasakawa's treatment of the same situation in a special case.

The usual definition of the optical-model scattering amplitude[9] is that it is the average of the actual scattering amplitude over an energy interval δ. In the actual situation, amplitudes for different energies are independent and it is the cross sections that are experimentally averaged. The meaning of the energy averaging process will be discussed.

In Sec. 5 we will discuss the possibility of performing scattering experiments with time definition which are complementary to the usual ones in which energy but not time is resolved. One such experiment has been suggested by the present authors.[10]

2. DEPENDENCE OF THE SCATTERING CROSS SECTION ON THE INITIAL BEAM CHARACTERISTICS

A beam consists of particles whose energy is defined within certain limits. The time of arrival of particles at a specific point may also be defined within limits by modulating the amplitude of the beam. In general the limits on the time and energy resolution of the beam are wider than those imposed by the uncertainty principle. For example some particles may be moving faster than others because they were accelerated a little more in the accelerator. The fast and slow beams could in principle be separated by a magnetic spectrometer. This type of energy uncertainty will be called classical uncertainty. On the other hand, if the time of arrival of a particle at a point is defined within a time interval τ, we cannot in principle measure its energy with a definition better than $\delta=\hbar/\tau$. This type of uncertainty will be called quantal uncertainty.

A beam with both classical and quantal uncertainty is described by a collection of wave packets with different mean wave numbers $\mathbf{k}_j$ whose position at time t is measured with respect to different initial position vectors $\mathbf{r}_i$. The value of $\mathbf{r}_i$ specifies the starting time of the wave packet at the source. One such wave packet may be written

$$\xi_{ij}(\mathbf{r},t)=(2\pi)^{-\frac{1}{2}}\int\Delta(\mathbf{k}_j,\delta;\mathbf{k})\exp[i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_i)]d^3k. \quad (2)$$

Suppose that the number of packets of type j is given by n where

$$n=\Phi_j N. \quad (3)$$

N is the total number of wave packets.

We wish to know the cross section for a scattering experiment with an incident beam which can be described in this way. For simplicity we will consider only one spin state. We must be careful to distinguish two cases.

Case A. If the $\mathbf{r}_i$ are such that the wave packets do not overlap both before and after scattering, the scattering of each wave packet may be considered as an individual event. The total cross section is the sum of the cross sections for the scattering of each wave packet.

Case B. The $\mathbf{r}_i$ are such that the wave packets overlap. In this case different wave packets will interfere and the cross sections for individual wave packets do not add incoherently.

The mathematical description of these cases is as follows. For case A we will consider wave packets with the same value $\mathbf{r}_0$ of $\mathbf{r}_i$ and describe the beam by means of a density matrix. The density matrix at time t is

$$\rho(\mathbf{r},\mathbf{r}',t)=\sum_j\Phi_j\xi_{0j}(\mathbf{r},t)\xi_{0j}^*(\mathbf{r}',t). \quad (4)$$

According to the standard theory of wave-packet scattering,[11] ξ_{0j} is given in the case where the space occupied by the wave packet is large compared to the size of the scatterer but small compared to the distance of the detector by

$$\xi_{0j}(\mathbf{r},t)=(2\pi)^{-\frac{1}{2}}\int\Delta(\mathbf{k}_j,\delta;\mathbf{k})\exp(-i\mathbf{k}\cdot\mathbf{r}_0-iEt/\hbar)\times[\exp(ikr)/r]f(k,\Omega)d^3k. \quad (5)$$

[7] I. E. McCarthy, *Proceedings of the International Symposium on Direct Interactions and Nuclear Reaction Mechanisms, Padua, 1962* (Gordon and Breach Publishers, Inc., New York, 1963).

[8] T. Sasakawa, Progr. Theoret. Phys. Suppl. 11, 69 (1959).

[9] For example G. E. Brown, Rev. Mod. Phys. 31, 893 (1959).

[10] L. R. Dodd and I. E. McCarthy, Phys. Rev. Letters 12, 136 (1964).

[11] See, for example, E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961). A review of general wave-packet scattering with applications to nuclear reactions has been given by N. Austern in *Selected Topics in Nuclear Theory* (International Atomic Energy Agency, Vienna, 1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$f(k,\Omega)$ is the scattering amplitude for scattering of particles into a detector whose angular position is Ω. $\Delta(\mathbf{k}_j,\delta;\mathbf{k})$ is a weight function which confines the wave packet to a volume in momentum space centered at $\mathbf{k}_j$ and of width δ. δ is the *quantal uncertainty*.

We will not be interested in the angular uncertainty in the wave packet, so we will write

$$\Delta(\mathbf{k}_j,\delta;\mathbf{k})=\frac{\delta(\theta_k-\theta_0)\delta(\phi_k-\phi_0)}{2\pi k^2\sin\theta_k}\Delta(k_j,\delta;k), \quad (6)$$

where (θ_k,ϕ_k), (θ_0,ϕ_0) are, respectively, the angular coordinates of $\mathbf{k}$ and $\mathbf{k}_j$, the beam direction.

The probability $P_j(t)$ of detecting a particle at time t at the detector whose position is $\mathbf{r}_d$ is, for the wave packet j,

$$P_j(t)=N_j|\xi_{0j}(\mathbf{r}_d,t)|^2r_d^2d\Omega. \quad (7)$$

This defines the normalization N_j of ξ_{0j}. The differential cross section is

$$(d\sigma/d\Omega)(\mathbf{k}_j,\delta)=v_0N_jr_d^2\int_{-\infty}^{\infty}|\xi_{0j}(\mathbf{r}_d,t)|^2dt, \quad (8)$$

where v_0 is the group velocity of the wave packet.

We may write the differential cross section for the beam described by the density matrix (4) as

$$d\sigma/d\Omega=v_0N_jr_d^2\int_{-\infty}^{\infty}\rho(\mathbf{r}_d,\mathbf{r}_d,t)dt. \quad (9)$$

Substituting from (4) and interchanging the order of summation and integration we have

$$d\sigma/d\Omega=\sum_j\Phi_jv_0N_jr_d^2\int_{-\infty}^{\infty}|\xi_{0j}(\mathbf{r}_d,t)|^2dt$$

$$=\sum_j\Phi_jd\sigma(\mathbf{k}_j,\delta)/d\Omega. \quad (10)$$

This is the result stated above for case A. Since we are only interested in the differences of the peak energies E_j for different wave packets j, we may replace the sum in (10) by an energy integral.

$$\frac{d\sigma}{d\Omega}=\int\Phi(E')\frac{d\sigma}{d\Omega}(E',\delta)dE'. \quad (11)$$

The cross section is the energy average of the cross sections for individual wave packets with different peak energies. In particular, for beams with negligible quantal uncertainty the cross section is the energy average of the cross sections for scattering events with plane wave and outgoing spherical wave boundary conditions. For this reason we will call ϕ, the width of the weight function $\Phi(E)$, the *classical uncertainty*.

We will now consider the more general case B and see under what circumstances the distinction between classical and quantal uncertainty can be maintained.

In terms of plane-wave states $\eta(\mathbf{k})$, the most general density matrix is

$$\rho(\mathbf{r},\mathbf{r}',t)=\int\int a(\mathbf{k},\mathbf{k}')\eta(\mathbf{k}')^*\eta(\mathbf{k})d^3kd^3k'. \quad (12)$$

We wish to write (12) in diagonal form in terms of wave-packet states ξ_{ij}.

$$\rho(\mathbf{r},\mathbf{r}',t)=\sum_{ij}\Phi_j\xi_{ij}\xi_{ij}^*. \quad (13)$$

We must know if this diagonalization is unique. If so, we can assign a definite classical weight factor Φ with uncertainty ϕ and a definite quantal weight factor Δ with uncertainty δ. If not, the distinction is meaningless.

In mathematical terms the question is as follows. Can the Hermitian matrix $\rho(\mathbf{r},\mathbf{r}',t)$ in (12) be diagonalized by more than one transformation of the type (2) of the basis vectors? A necessary condition for the diagonalization to be unique[12] is that the states ξ_{ij} must be orthogonal like the states η.

In fact we have

$$\int\xi_{ij}\xi_{i'j'}d^3r=(2\pi)^{-3}\int\Delta(\mathbf{k}_j,\delta;\mathbf{k})\Delta(\mathbf{k}_{j'},\delta;\mathbf{k})^*$$

$$\times\exp[-i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_{i'})]d^3k. \quad (14)$$

The overlap integral in (14) is approximately zero if the wave packets have sufficiently different starting times given by $\mathbf{r}_i$ and $\mathbf{r}_{i'}$ because of the rapid oscillations of the integrand.

In case A the wave packets have a definite physical significance. The quantal uncertainty leads to alterations in the usual cross sections, as will be shown in the succeeding sections. In case B the wave packets may be regarded as mathematical aids to the visualization of the scattering process. For example we can decompose a plane wave into an infinite number of overlapping wave packets and then follow the propagation of one of these packets. Nevertheless, this would be without physical significance because the scattering amplitudes for all the packets must be recombined to obtain the cross section.

The vital point is that the wave packets both before and after the scattering must be quite distinct in space and time for the quantal uncertainty to have any physical significance.

3. SCATTERING OF A WAVE PACKET FROM A RESONANCE

In the usual theory of scattering it is assumed that $f(k,\Omega)$ in Eq. (5) varies much more slowly with energy than the wave packet amplitude factor $\Delta(\mathbf{k}_j,\delta;\mathbf{k})$; that is, that the quantal uncertainty δ is negligible in com-

[12] U. Fano, Rev. Mod. Phys. 29, 74 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

parison with the width of the energy fluctuations of the scattering amplitude.

We will consider the case where δ is comparable with Γ_s, the width of the compound state $|s\rangle$ of the system comprising the incident particle and the scatterer whose energy ϵ_s is near the central energy E_0 of the wave packet. In the present section we will suppose that only one state $|s\rangle$ contributes to the scattering amplitude at the energies covered by the wave packet.[12a] We will be interested in only one wave packet $j=0$.

Since we are interested in the energy rather than the momentum of the wave packet, it will be convenient to transform the integral in (5) to an energy integral. The transformation is trivial when k is proportional to E as it is for photons ($E=\hbar ck$).

For our discussion of hypothetical time-dependent experiments with finite-mass particles complementary to the usual energy-dependent experiments, we can make the assumption, following Friedman and Weisskopf,[4] that the source and detector are sufficiently close to the scatterer for wave packet spreading to be negligible. We will neglect the final term in the following expansion of E:

$$E=\frac{\hbar^2}{2m}[2\mathbf{k}\cdot\mathbf{k}_0-k_0^2+|\mathbf{k}-\mathbf{k}_0|^2]. \tag{15}$$

Photon wave packets do not spread for practical purposes.

We must now consider a particular form for the energy amplitude factor $\Delta(E_0,\delta;E)$ corresponding to $\Delta(\mathbf{k}_0,\delta;\mathbf{k})$. Wave packets made by electronic means (see Sec. 5) would have a rise and decay function in time that is something like an exponential. We will therefore consider a wave packet that has an exponential rise and decay, with time constant $\hbar/\delta$. We will assume that the beam is switched off as soon as it has attained full strength, so that there is no time for which the beam intensity is constant. The experimental difficulty is in fast switching on and off. Since we want as short a packet in time as possible, this is then the most realistic form. The energy amplitude factor for this form is

$$\Delta(E_0,\delta;E)=\frac{\delta/2\pi}{(E-E_0)^2+\delta^2/4}. \tag{16}$$

One shape of wave packet is already available in nature. This is the photon wave packet from the decay of a level $|s\rangle$ of a many-body system. It has a half-exponential time spectrum. If the moment of excitation of the level is taken as zero time, then the probability of emission of a photon decays exponentially with a time constant $\hbar/\Gamma_s$. In this case we have

$$\Delta(E_0,\Gamma_s;E)=\frac{i/2\pi}{E-E_0+i\Gamma_s/2}. \tag{17}$$

We will use the form (16) for most of our discussion.

For the scattering amplitude we will use the expansion of Siegert, Humblet, and Rosenfeld,[13]

$$f[k(E),\theta]=\frac{1}{2i}\sum_l(2l+1)P_l(\cos\theta)\times\left[C_l(E)+\sum_n\frac{R_{ln}}{E-\epsilon_{ln}+\frac{1}{2}i\Gamma_{ln}}\right]. \tag{18}$$

The nonresonant term $C_l(E)$ varies slowly with energy. The splitting off of this term is not unique and can be made according to some physical prescription. Equation (5) now becomes

$$\xi(\mathbf{r},t)=\frac{K}{2\pi}\sum_l(2l+1)P_l(\cos\theta)\int_0^\infty\frac{\delta}{(E-E_\delta)(E-\bar{E}_\delta)}\times\exp(iEX)\left[C_l(E)+\sum_n\frac{R_{ln}}{E-\epsilon_{ln}+i\Gamma_{ln}/2}\right]dE, \tag{19}$$

where

$$X=[(r_0+r)/v_0-t]/\hbar,\quad E_\delta=E_0-i\delta/2,\quad K\bar{K}=1/8\pi r^2. \tag{20}$$

We have dropped the subscripts on $\xi(\mathbf{r},t)$ because we are now interested only in one particular wave packet.

The integral in (19) is now in a convenient form for contour integration if we make the approximation of extending the lower limit to $-\infty$, thus neglecting the contributions of bound states.

We will first consider the scattering of the wave packet according to the nonresonant scattering amplitude $C_l(E)$ whose variation with energy can be neglected over the energies of the wave packet. Performing the integration in Eq. (19) we find

$$\begin{aligned}\xi(\mathbf{r},t)&=K\sum_l(2l+1)P_l(\cos\theta)C_l(E)\exp(i\bar{E}_\delta X),\quad X>0\\&=K\sum_l(2l+1)P_l(\cos\theta)C_l(E)\times\exp(iE_\delta X),\quad X<0.\end{aligned} \tag{21}$$

The wave packet is centered at $X=0$, i.e., at

$$t=(r_0+r)/v_0. \tag{22}$$

Hence, the nonresonant packet is propagated without time delay. The time spectrum of $|\xi(\mathbf{r},t)|^2$ is an exponential rise and fall, $\exp(-\delta t/\hbar)$.

The differential cross section is, according to Eq. (8),

$$d\sigma/d\Omega=\tfrac{1}{4}|\textstyle\sum_l(2l+1)P_l(\cos\theta)C_l(E)|^2. \tag{23}$$

Thus, the differential cross section for nonresonant (potential) scattering is independent of the quantal uncertainty δ and identical with that for a normal beam in which δ is negligible.

[12a] *Note added in proof*. The decay of a single resonance has been investigated by R. G. Newton, Ann. Phys. (N. Y.) 14, 333 (1961).

[13] A. Siegert, Phys. Rev. 56, 750 (1939). J. Humblet and L. Rosenfeld, Nucl. Phys. 26, 529 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We will now consider the scattering from an isolated resonance state $|s\rangle$. Taking one term of the sum in (18) we have for the integral in (18)

$$I=\int_{-\infty}^{\infty}\frac{\delta/2\pi}{(E-E_\delta)(E-\bar{E}_\delta)}\frac{R_s}{E-E_s}\exp(iEX)dE, \quad (24)$$

where

$$E_s=\epsilon_s-i\Gamma_s/2. \quad (25)$$

Integration round an infinite semicircle in the upper half plane gives

$$I=\frac{R_s}{\bar{E}_\delta-E_s}\exp(i\bar{E}_\delta X), \quad X>0. \quad (26)$$

This means that the leading edge of the wave packet, that is, the part for times greater than $(r_0+r)/v_0$, is propagated with the same shape as it originally had, but, of course, with a different magnitude. The propagation of the trailing edge (the tail) of the wave packet is given by integration round a contour in the lower half plane

$$I=\frac{R_s}{E_\delta-E_s}\exp(iE_\delta X)$$
$$-i\delta\frac{R_s}{(E_s-E_\delta)(E_s-\bar{E}_\delta)}\exp(iE_sX), \quad X<0. \quad (27)$$

The time spectrum given by $|I|^2$ is a decaying oscillatory function. In particular, if the time width of the wave packet is much less than the decay constant of the state $|s\rangle$, the time spectrum of the tail has the shape $\exp(-\Gamma_s t/\hbar)$. This corresponds to exciting the resonance suddenly and watching it decay with its natural time constant.

The differential cross section is

$$\frac{d\sigma}{d\Omega}=\frac{R_s\bar{R}_s[(E-\epsilon_s)^2\Gamma_s/2+(\Gamma_s+\delta)^2(\Gamma_s/2+\delta)/4]}{2\Gamma_s[(E-\epsilon_s)^2+(\Gamma_s+\delta)^2/4]^2}. \quad (28)$$

This reduces to the Breit-Wigner form for very small quantal uncertainty δ. The most significant thing is that the width of the energy spectrum is $\Gamma_s+\delta$. Note also that the differential cross section for scattering from a single resonance is reduced in magnitude for large quantal uncertainty. For large δ, $d\sigma/d\Omega$ is of order $1/\delta$.

In view of the arguments often made about the noninterference of potential and resonance scattering for large δ, it is interesting to consider the interference term. Denote the additional cross section due to the interference of potential and resonant scattering by $d\sigma_I/d\Omega$. It arises from the last term in the splitting of the scattering into potential and resonant parts

$$|\xi(\mathbf{r},t)|^2=|\xi_P(\mathbf{r},t)|^2+|\xi_R(\mathbf{r},t)|^2$$
$$+2\,\mathrm{Re}\bar{\xi}_P(\mathbf{r},t)\xi_R(\mathbf{r},t). \quad (29)$$

Taking only the s state for simplicity of notation,

$$\frac{d\sigma_I}{d\Omega}=\tfrac{1}{2}\left\{R_s\bar{C}_0(E)\frac{(E-\epsilon_s)+i(\Gamma_s/2+\delta)}{[(E-\epsilon_s)+i(\Gamma_s+\delta)^2/4]^2}\right\}, \quad (30)$$

$$=\tfrac{1}{2}\,\mathrm{Re}[R_s\bar{C}_0(E)]\frac{(E-\epsilon_s)(E-\epsilon_s)^2+(\Gamma_s+\delta)(\Gamma_s+3\delta)/4}{[(E-\epsilon_s)^2+(\Gamma_s+\delta)^2/4]^2}$$
$$+\tfrac{1}{2}\,\mathrm{Im}[R_s\bar{C}_0(E)]\frac{(E-\epsilon_s)^2\Gamma_s/2+(\Gamma_s+\delta)^2(\Gamma_s+2\delta)/8}{(E-\epsilon_s)^2+(\Gamma_s+\delta)^2/4}. \quad (31)$$

The coefficient of $\mathrm{Im}[R_s\bar{C}_0(E)]$ is of order $1/\delta$ for large δ. Hence the interference term in the cross section does become smaller for larger δ, that is for better defined wave packets in time, but so also does the resonant scattering term, both being of order $1/\delta$. The potential scattering term is not affected by δ.

4. SCATTERING FROM MANY LEVELS AND THE OPTICAL MODEL

We will now consider the energy region where, for a given channel, the resonances in the scattering amplitude are such that the average width is greater than the average spacing, so that many levels contribute at each energy over the energy spread of the wave packet.

The levels in this region are defined by some plausible model such as has been discussed by Brown[9] for nuclei. Their widths are of the order of 1 eV, so that the corresponding lifetime, 10^{-15} sec, is too short for a practical wave-packet experiment. However, it is interesting to consider a hypothetical wave-packet experiment in order to give a proper quantum mechanical sequential description of the passage of a particle through the scatterer, when we know the S-matrix element for the scattering.

This gives us another way of looking at the optical and direct interaction models. We will restrict ourselves to elastic scattering for simplicity. The optical model has been considered in two ways, each with a different

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

starting point. The first way starts with a Schrödinger equation. The optical-model Hamiltonian is shown to be an approximation to the many-body Hamiltonian. The second way is independent of any Hamiltonian assumption and starts with the scattering amplitude. The average over some energy interval I of the scattering amplitude S is shown to be capable of being calculated from an optical-model Hamiltonian. The average cross section is split into two terms.

$$\langle\sigma\rangle=|\langle S\rangle|^2+\langle|S-\langle S\rangle|^2\rangle. \tag{32}$$

The first term is the optical-model or shape elastic cross section. The second term is the fluctuation or compound elastic cross section.

A third, classical, definition of the optical model is sometimes given. The optical-model cross section is the cross section for propagation without time delay while the compound elastic cross section is the remainder.

It is not clear that the last two definitions are equivalent. The splitting in Eq. (32) requires a definition of the averaging procedure. Some authors, for example Brown,[9] have gone to much trouble to discuss the averaging procedure. Brown's argument is essentially a wave-packet argument since it uses an average of the scattering amplitude over an energy interval I. We have seen in Sec. 2 that for a beam without significant time resolution it is the cross sections, not the amplitudes, that are averaged.

Another condition for the validity of this definition of the optical model is required. The average over amplitudes must be nearly equivalent to the average over cross sections. This means that the compound elastic cross section must be very small compared with the optical model cross section.

A quantal statement of the third definition of the optical model must be given by a detailed wave-packet description. It is quite conceivable that the scattering amplitudes for certain problems are such that a large proportion of the cross section is due to time-delayed wave packets. This question will be discussed in detail by one of us (LRD) in a subsequent publication.

At present we will just consider the propagation of the wave packet in general and show how it leads to a large proportion of immediate propagation in the case $\Gamma\gg D$ where Γ is the average level width and D is the average spacing.

Considering only the S wave for simplicity and omitting the potential scattering term, Eq. (19) gives for the time dependence of the trailing edge of the wave packet

$$\xi(\mathbf{r},t)=K\exp(iEX)\sum_n\frac{R_n}{E-\epsilon_n+i(\Gamma_n-\delta)/2}\Bigg\{\exp(\delta X/2)-\frac{i\delta\exp(\Gamma_n X/2)\exp i(\epsilon_n-E)X}{E-\epsilon_n+i(\Gamma_n+\delta)/2}\Bigg\}. \tag{33}$$

The second term in the bracket is the one containing the properties of the scattering amplitude rather than the wave packet. Each level n contributes an exponential tail to the amplitude with a time-delay constant $\hbar/\Gamma_n$. However, this contribution is multiplied by a phase factor $\exp[i(\epsilon_n-E)X]$ which gives a partial cancellation of the tails when the packets from different levels are superposed.

If the level widths Γ_n are much greater than the average spacing D, the factor ϵ_n-E can be large so that the phase factor oscillates rapidly. In this case the phases of the contributions from different levels tend to be random so the tails cancel out giving a large proportion of propagation without time delay.

Thus the sequential description of scattering gives the same result as the usual energy description. The optical model is valid when $\Gamma\gg D$.

The dependence of (33) on the magnitude of the quantal uncertainty δ is also interesting. As δ is increased the magnitude of the contribution from each level to the scattered packet decreases. However more levels contribute significantly to the sum. If the phases of the residues R_n are random the magnitude of the scattered packet will decrease rapidly with increasing δ. If they are correlated, the magnitude will decrease less rapidly. The range of the correlations between levels can be determined in principle by varying δ. This will be discussed in detail in a subsequent publication.

5. POSSIBLE WAVE-PACKET EXPERIMENTS

The condition for a time-dependent scattering experiment is that the experimental definition of time must be accurate in comparison with the characteristic time of the scattering amplitude. That is, we must have δ not much less than Γ.

Experimental definition of time is at present possible for times as short as about 10^{-10} sec. Typical nuclear values for $\hbar/\Gamma$ are 10^{-15} sec, so wave-packet experiments cannot be performed with nuclei except in special cases. These cases are metastable states which can have lifetimes as long as 10^{-7} sec.

One experiment has actually been done using the Mössbauer effect with the 14-keV γ ray from Fe^{57}, which has a decay constant of 10^{-7} sec, by Holland, Lynch, Perlow, and Hanna.[14] The time spectrum of the incident wave packet was defined by using as zero time the time of formation of the 14-keV state, which was defined by the time of emission of the 128-keV γ ray (a fast decay) from the next highest state in the γ-ray cascade from the decay of Co^{57}. The wave packet has an exponential time spectrum with $\delta=\hbar/10^{-7}$ sec. This wave packet was scattered resonantly from an Fe^{57} target. The time spectrum of the final state and the increased width of the absorption line were both

[14] R. E. Holland, F. J. Lynch, G. J. Perlow, and S. S. Hanna, Phys. Rev. Letters 4, 181 (1960); F. J. Lynch, R. E. Holland, and M. Hamermesh, Phys. Rev. 120, 513 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

observed. In this case $\delta=\Gamma_s$. The width of the wave packet in this experiment is of course fixed. By looking at the scattering at times less than 10^{-7} sec, Holland *et al.* were able to observe greater widths, but for these cases the shape of the incident packet was not defined. The spectrum of total elapsed time gives only an upper limit to the time width of the incident wave packet because it is not known if the delay occurred in the source or the scatterer.

The Mössbauer effect is, of course, a wave-packet scattering experiment in which the absorption cross section is measured. If we do not define time but merely observe the cross section we lose the wave-packet property, that is the quantal uncertainty. The cross section is obtained using Eq. (17) to define the energy amplitude factor of the wave packet.

$$\frac{d\sigma}{d\Omega}=\frac{R_s^2}{2\Gamma_s}\frac{1}{(E_0-E_s)^2+\Gamma_s^2}. \tag{34}$$

Thus, we have the well-known result that the linewidth in the Mössbauer effect is twice the width of the level, assuming all the nuclei in the target are capable of absorbing resonantly.

Defining the starting time of the excitation of the metastable state puts a quantal uncertainty into the beam equal to Γ_s. The method of Holland *et al.* defines the starting time with a minimum tolerance equal to the lifetime of the next highest state in the γ-ray cascade. This is shorter than the experimental time resolution, which is itself much shorter than $\hbar/\Gamma_s$.

One way of varying the time width of the wave packet would be to vary the resolution with which the starting time is measured. It is possible by this method to decrease δ, but not to increase it.

A better experimental way of varying the wave-packet width in the Mössbauer effect has been suggested by the present authors. A third resonant absorber is introduced between the source and the target. This absorber is accelerated in a very short time (10^{-9} sec) to a speed sufficient to shift the resonance so that the incident beam is no longer absorbed and can hit the target. It is then slowed down again quickly so that the time duration of the pulse is of the order of 10^{-9} sec. This method would produce approximately the exponential wave packet [Eq. (16)] that we have used in the calculations of Secs. 3 and 4. It is experimentally just possible to achieve the requisite acceleration by using a very thin foil of Fe^{57} as one plate of a parallel plate condenser in a vacuum which is charged first with one sign and then with the opposite sign by an rf pulse. However, the acceleration may be achieved more easily using a piezo-electric crystal. Another possibility is to use the Stark effect to shift the resonance. This requires a tightly bound dielectric crystal containing nuclei with a metastable state.

Wave-packet experiments in the atomic energy region may be interesting. An absorber of laser material placed in a laser beam and moved for a short time as suggested above would produce wave packets of laser intensity. The quantal uncertainty would be much greater than that of a beam from a pulsed laser, whose quantal uncertainty can be no greater than the width of the laser state.

All such experiments observe only the scattering of a wave packet from a single resonance with a trivial angular distribution. Overlapping resonances, even in atoms, would probably have widths of the order of a few tenths of an electron volt. Wave-packet experiments in this region with $\delta\sim\Gamma$ would require time lengths of 10^{-15} sec which seems impossible at present. The possibility of doing time-dependent electron-scattering experiments from atoms is not experimentally remote and would be interesting.

Using laser wave packets, it may be possible to observe the changes in angular distribution due to the interference of potential and resonant scattering as the quantal uncertainty is changed. The potential scattering could be obtained by diffraction of laser light round a small crystal of laser material.

ACKNOWLEDGMENTS

We would like to acknowledge very helpful criticism from Dr. C. A. Hurst and helpful discussions with Professor W. J. Knox and Professor K. M. Watson.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Coherence Properties of Blackbody Radiation.* I. Correlation Tensors of the Classical Field

C. L. MEHTA AND E. WOLF
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 9 December 1963)

This paper is concerned with the extension of some recently reported results, especially those of Bourret (1960), relating to coherence properties of blackbody radiation. An explicit expression for the complex electric correlation tensor $\mathcal{E}_{ij}(\mathbf{r},\tau)$ of blackbody radiation is derived, and on the basis of it spatial coherence ($\tau=0$) is discussed in detail. The behavior of diagonal as well as nondiagonal components is illustrated by contour diagrams. In particular, it is found that the nondiagonal components of the correlation tensor, even though being zero for zero-space separation ($\mathbf{r}=0$), acquire, in general, nonvanishing values when $\mathbf{r}\neq 0$. The magnetic and mixed correlation tensors are also discussed.

1. INTRODUCTION

IT is a common notion that the most incoherent radiation is blackbody radiation in an equilibrium enclosure. Recent researches[1–4] have shown that even in this type of radiation there is coherence in a sufficiently small space-time region. Bourret[1] has derived expressions for the second-order electric correlation tensor of blackbody radiation, using techniques analogous to those employed in the theory of isotropic turbulence of an incompressible fluid. Such a correlation tensor describes the correlation between the Cartesian components (denoted by subscripts i, j) of the electric field $\mathbf{E}^{(r)}(\mathbf{r},t)$ at two space-time points $\mathbf{r}_1$ and $\mathbf{r}_2$, at time instants t_1 and t_2,

$$\mathcal{E}_{ij}^{(r)}(\mathbf{r}_1,\mathbf{r}_2,t_1,t_2)=\langle E_i^{(r)}(\mathbf{r}_1,t_1)E_j^{(r)}(\mathbf{r}_2,t_2)\rangle, \quad (1.1)$$

where sharp brackets denote the time average:

$$\langle E_i^{(r)}(\mathbf{r}_1,t_1)E_j^{(r)}(\mathbf{r}_2,t_2)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} E_i^{(r)}(\mathbf{r}_1,t_1+t)E_j^{(r)}(\mathbf{r}_2,t_2+t)dt. \quad (1.2)$$

Sarfatt[3] rederived some of Bourret's results using quantum mechanical density matrix techniques.

The components of the correlation tensors discussed by Bourret and Sarfatt are real functions of space and time. However, numerous recent researches on coherence properties of light have shown that an appropriate measure of coherence is provided by certain complex rather than real correlation functions (cf. Ref. 5), and this is also supported by recent investigations of Glauber[6] based on quantum field theoretical calculations. (See also Sudarshan,[7a,b] Wolf,[8] Mandel, Sudarshan and Wolf,[9] Mandel.[10]) The appropriate complex electric correlation tensor $\mathcal{E}_{ij}(\mathbf{r}_1,\mathbf{r}_2,t_1,t_2)$ may be derived from the real correlation tensor defined by (1.1), by using the concept of an analytic signal.[11,5] Assuming the field to be stationary in time, so that $\mathcal{E}_{ij}^{(r)}$ depends on t_1 and t_2 only through the time difference $\tau=t_1-t_2$, the complex correlation tensor is given by[12]

$$\mathcal{E}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=2[\mathcal{E}_{ij}^{(r)}(\mathbf{r}_1,\mathbf{r}_2,\tau)+i\mathcal{E}_{ij}^{(i)}(\mathbf{r}_1,\mathbf{r}_2,\tau)], \quad (1.3)$$

where

$$\mathcal{E}_{ij}^{(i)}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{\mathcal{E}_{ij}^{(r)}(\mathbf{r}_1,\mathbf{r}_2,\tau')}{\tau'-\tau}d\tau' \quad (1.4)$$

is the Hilbert transform of $\mathcal{E}_{ij}^{(r)}(\mathbf{r}_1,\mathbf{r}_2,\tau)$ and P denotes the Cauchy principal value at $\tau'=\tau$.

Bourret and Sarfatt have restricted their discussion of the tensor to certain special cases only, namely, to those characterizing temporal correlation ($\mathbf{r}_1=\mathbf{r}_2$) and lateral and longitudinal spatial correlation ($\tau=0$). Their results do not provide information about the nondiagonal components of the correlation tensor, though, in principle such information could be obtained with the help of their formulas.

* This research was supported by the U. S. Air Force Office of Scientific Research.

[1] R. C. Bourret, Nuovo Cimento 18, 347 (1960).
[2] Y. Kano and E. Wolf, Proc. Phys. Soc. (London) 80, 1273 (1962).
[3] J. Sarfatt, Nuovo Cimento 27, 1119 (1963).
[4] C. L. Mehta, Nuovo Cimento 28, 401 (1963).
[5] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959), Chap. X.
[6] R. J. Glauber, Phys. Rev. 130, 2529 (1963).
[7] E. C. G. Sudarshan, (a) Phys. Rev. Letters 10, 277 (1963); (b) in *Proceedings of the Symposium on Optical Masers* (Polytechnic Press, Brooklyn, New York and J. Wiley & Sons, Inc., New York, 1963), p. 45.
[8] E. Wolf, in *Proceedings of the Symposium on Optical Masers* (Polytechnic Press, Brooklyn, New York, and John Wiley & Sons Inc., New York, 1963), p. 29.
[9] L. Mandel, E. C. G. Sudarshan, and E. Wolf (to be published).
[10] L. Mandel, Phys. Letters 7, 117 (1963).
[11] D. Gabor, J. Inst. Electr. Engrs. 93, Part III, 429 (1946).
[12] An alternative, but equivalent, definition of the complex correlation tensor, which will be needed later is as follows: With the real field $\mathbf{E}^{(r)}(\mathbf{r},t)$ we associate the complex analytic signal $\mathbf{E}(\mathbf{r},t)$, i.e.,

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}^{(r)}(\mathbf{r},t)+i\mathbf{E}^{(i)}(\mathbf{r},t),$$

where $\mathbf{E}^{(i)}$ is the Hilbert transform of $\mathbf{E}^{(r)}$. Then the complex correlation tensor $\mathcal{E}_{ij}$ may be expressed in the form

$$\mathcal{E}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\langle E_i(\mathbf{r}_1,t+\tau)E_j^*(\mathbf{r}_2,t)\rangle. \quad (1.3a)$$

The equivalence of the definitions (1.3) and (1.3a) is shown in Ref. 17, pp. 464–466, where also a certain mathematical refinement connected with questions of convergence is discussed.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In this paper an explicit expression for the complete second-order complex electric correlation tensor of blackbody radiation is derived and a number of diagrams are obtained, which show the behavior of the elements of this tensor. The magnetic and the mixed second-order complex correlation tensors are also discussed.

In paper II (Ref. 13) the electromagnetic correlation tensors defined on the basis of quantum field theory will be evaluated for blackbody radiation and will be compared with the corresponding tensors defined in the present paper on the basis of semiclassical theory. Higher order correlation tensors will also be considered in paper II.

2. EXPLICIT EXPRESSION FOR THE SECOND-ORDER COMPLEX ELECTRIC CORRELATION TENSOR OF BLACKBODY RADIATION

In this section a new expression for the normalized complex electric correlation tensor $\gamma_{ij}(\mathbf{r},\tau)$ of blackbody radiation will be derived. For any stationary field, the normalized correlation tensor is defined by the formula

$$\gamma_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\frac{\mathcal{E}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)}{[\mathcal{E}_{ii}(\mathbf{r}_1,\mathbf{r}_1,0)]^{1/2}[\mathcal{E}_{jj}(\mathbf{r}_2,\mathbf{r}_2,0)]^{1/2}}, \quad (2.1)$$

where $\mathcal{E}_{ij}$ is the correlation function (1.3). The normalization ensures[5] that $0\leq|\gamma_{ij}|\leq 1$. In the case of blackbody radiation in a cavity whose linear dimensions are large compared with the mean wavelength of the radiation, the field is isotropic and $\mathcal{E}_{ij}$ and γ_{ij} depend effectively on $\mathbf{r}_1$ and $\mathbf{r}_2$ through the difference $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ only.

An integral expression for $\gamma_{ij}(\mathbf{r},\tau)$ for blackbody radiation has been obtained by Kano and Wolf.[2] They showed that

$$\gamma_{ij}(\mathbf{r},\tau)=\frac{45\alpha^4}{8\pi^5}\int\frac{k^2\delta_{ij}-k_ik_j}{k\{\exp(\alpha k)-1\}}\times\exp\{i(\mathbf{k}\cdot\mathbf{r}-kc\tau)\}d^3k, \quad (2.2)$$

where

$$\alpha=\hbar c/KT, \quad (2.3)$$

c being the vacuum velocity of light, $\hbar$ the Planck's constant divided by 2π, K the Boltzmann constant and T the absolute temperature. The integration in (2.2) is taken over the whole $\mathbf{k}$ space. The normalization constant in (2.1) has the value

$$[\mathcal{E}_{ii}(\mathbf{r}_1,\mathbf{r}_1,0)]^{1/2}[\mathcal{E}_{jj}(\mathbf{r}_1,\mathbf{r}_2,0)]^{1/2}\equiv\mathcal{E}_{ii}(0,0,0) \quad \text{(no summation)}$$

$$=\frac{64}{45}\frac{\pi^6K^4T^4}{(hc)^3}. \quad (2.4)$$

Equation (2.2) may be written as

$$\gamma_{ij}(\mathbf{r},\tau)=\frac{45\alpha^4}{8\pi^5}(\partial_i\partial_j-\delta_{ij}\nabla^2)\int\frac{\exp\{i(\mathbf{k}\cdot\mathbf{r}-kc\tau)\}}{k(e^{\alpha k}-1)}d^3k, \quad (2.5)$$

where $\partial_i\equiv\partial/\partial r_i$. Using spherical polar coordinates for $\mathbf{k}$, with the polar axis along the direction of $\mathbf{r}$, we obtain

$$\gamma_{ij}(\mathbf{r},\tau)=\frac{45\alpha^4}{8\pi^5}(\partial_i\partial_j-\delta_{ij}\nabla^2)\int_0^\infty kdk\frac{e^{-ikc\tau}}{e^{\alpha k}-1}\int_0^\pi e^{ikr\cos\theta}\sin\theta d\theta\int_0^{2\pi}d\phi$$

$$=\frac{45\alpha^4}{2\pi^4}(\partial_i\partial_j-\delta_{ij}\nabla^2)\frac{1}{r}\int_0^\infty\frac{\sin kr\,e^{-ikc\tau}}{e^{\alpha k}-1}dk. \quad (2.6)$$

The last integral can be written as

$$\sum_{n=1}^\infty\frac{1}{2i}\int_0^\infty\{e^{ik(r-c\tau)}-e^{-ik(r+c\tau)}\}e^{-n\alpha k}dk=\frac{1}{2i}\sum_{n=1}^\infty\left\{\frac{1}{n\alpha-i(r-c\tau)}-\frac{1}{n\alpha+i(r+c\tau)}\right\},$$

so that γ_{ij} may be expressed in the form

$$\gamma_{ij}(\mathbf{r},\tau)=\frac{45\alpha^4}{2\pi^4}(\partial_i\partial_j-\delta_{ij}\nabla^2)\sum_{n=1}^\infty\frac{1}{(n\alpha+ic\tau)^2+r^2}.$$

Carrying out the differentiation on the right-hand side, we finally obtain the following expression for the normalized complex electric correlation tensor of blackbody radiation:

$$\gamma_{ij}(\mathbf{r},\tau)=\frac{90\alpha^4}{\pi^4}\sum_{n=1}^\infty\left\{\frac{\delta_{ij}}{\{(n\alpha+ic\tau)^2+r^2\}^2}+2\frac{r_ir_j-r^2\delta_{ij}}{\{(n\alpha+ic\tau)^2+r^2\}^3}\right\}, \quad (2.7)$$

where r_i, r_j are the components of the vector $\mathbf{r}$ with respect to i and j axes, respectively.

3. TEMPORAL COHERENCE

Restricting ourselves first to the case $\mathbf{r}=0$, Eq. (2.7) gives the following expression for the normalized tensor that characterizes *temporal coherence*:

$$\gamma_{ij}(0,\tau)=(90/\pi^4)\zeta(4,1+ic\tau/\alpha)\delta_{ij}, \quad (3.1)$$

where $\zeta(s,a)$ is the generalized Riemann zeta function[14]

[13] C. L. Mehta and E. Wolf, following paper, Phys. Rev. 134, A1149 (1964).

[14] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, New York, 1958), p. 265; (also Dover Publications, Inc., New York, 1962), p. 265.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

defined by

$$\zeta(s,a)=\sum_{n=0}^{\infty}\frac{1}{(n+a)^s}. \tag{3.2}$$

It should be noted that the tensor $\gamma_{ij}(0,\tau)$ is *diagonal.*

The expression (3.1) for the temporal degree of coherence of blackbody radiation is identical with the expression derived by Kano and Wolf[2] and discussed in detail by them [cf. Eq. (11) of their paper]. In particular they gave diagrams showing the behavior of the modulus and the argument of $\gamma_{ii}(0,\tau)$ as function of τ.

4. SPATIAL COHERENCE

Setting $\tau=0$ in Eq. (2.7) we obtain the following expression for the normalized tensor that characterizes *spatial coherence*:

$$\gamma_{ij}(\mathbf{r},0)=\frac{90}{\pi^4}\sum_{n=1}^{\infty}\left\{\frac{\delta_{ij}}{(n^2+r^2/\alpha^2)^2}+\frac{2}{\alpha^2}\frac{r_ir_j-\delta_{ij}r^2}{(n^2+r^2/\alpha^2)^3}\right\}. \tag{4.1}$$

It is seen from this expression that $\gamma_{ij}(\mathbf{r},0)$ is real. The series on the right-hand side of (4.1) may easily be summed with the help of the result[15]

$$\sum_{n=1}^{\infty}\frac{1}{n^2+a^2}=\frac{\pi}{2a}L(\pi a), \tag{4.2}$$

where $L(x)$ is the Langevin function

$$L(x)=\coth x-1/x. \tag{4.3}$$

Successive term by term differentiation of (4.2) gives

$$\sum_{n=1}^{\infty}\frac{1}{(n^2+a^2)^2}=-\frac{1}{2a}\frac{d}{da}\left\{\frac{\pi}{2a}L(\pi a)\right\}, \tag{4.4}$$

$$\sum_{n=1}^{\infty}\frac{1}{(n^2+a^2)^3}=\frac{\pi}{16a}\frac{d}{da}\left\{\frac{1}{a}\frac{d}{da}\left[\frac{1}{a}L(\pi a)\right]\right\}. \tag{4.5}$$

If we set

$$\mathbf{r}=(\alpha/\pi)\mathbf{r}', \tag{4.6}$$

so that r' represents the separation of points in the cavity in units of α/π, (4.1) gives, with the help of (4.4), (4.5), and (4.3),

$$\gamma_{ij}(\mathbf{r}')\equiv\gamma_{ij}(\mathbf{r},0)=\frac{45}{4r'^4}\left[A(r')\delta_{ij}+B(r')\frac{r_i'r_j'}{r'^2}\right], \tag{4.7}$$

where

$$\left.\begin{aligned}A(r')&=-r'\coth r'-r'^2\operatorname{csch}^2r'\\&\quad-2r'^3\operatorname{csch}^2r'\coth r'+4,\\B(r')&=3r'\coth r'+3r'^2\operatorname{csch}^2r'\\&\quad+2r'^3\operatorname{csch}^2r'\coth r'-8.\end{aligned}\right\} \tag{4.8}$$

[15] L. B. W. Jolley, *Summation of Series* (Dover Publications, New York, 1961), p. 22, Series number 124.

Equation (4.7) is valid for all $\mathbf{r}'$ but is not suitable for computing $\gamma_{ij}(\mathbf{r}')$ when r' is small. For small r', however, we directly obtain from (4.1), on expanding in powers of r',

$$\gamma_{ij}(\mathbf{r}')\equiv\gamma_{ij}(\mathbf{r},0)=\delta_{ij}\left[1-\frac{8}{21}r'^2+\frac{3}{35}r'^4+\cdots\right]+\frac{r_i'r_j'}{r'^2}\left[\frac{4}{21}r'^2-\frac{2}{35}r'^4+\cdots\right]\quad(r'<\pi). \tag{4.9}$$

If we set $i=j=x$ in (4.7) and (4.9) and take $\mathbf{r}'$ in the direction of the x axis, we obtain the following expressions for the *longitudinal coherence function* $\gamma_{xx}^{\text{long}}$:

$$\gamma_{xx}^{\text{long}}(\mathbf{r}')=\frac{45}{2r'^4}(r'\coth r'+r'^2\operatorname{csch}^2r'-2)\qquad(r'<\infty), \tag{4.10a}$$

$$=1-\frac{4}{21}r'^2+\frac{1}{35}r'^4+\cdots\qquad(r'<\pi), \tag{4.10b}$$

where $r'=x'=\pi x/\alpha$.

On the other hand, if $\mathbf{r}'$ is chosen perpendicular to the x axis, then we obtain from (4.7) the following expression for *lateral coherence function* γ_{xx}^{lat}:

$$\gamma_{xx}^{\text{lat}}(\mathbf{r}')=\frac{45}{4r'^4}[4-r'\coth r'-r'^2\operatorname{csch}^2r'-2r'^3\coth r'\operatorname{csch}^2r'],\qquad(r'<\infty), \tag{4.11a}$$

$$=1-\frac{8}{21}r'^2+\frac{3}{35}r'^4+\cdots\qquad(r'<\pi), \tag{4.11b}$$

where $r'=(y'^2+z'^2)^{1/2}=\pi(y^2+z^2)^{1/2}/\alpha$.

Expressions for $\gamma_{yy}^{\text{long}}$, γ_{yy}^{lat}, $\gamma_{zz}^{\text{long}}$, and γ_{zz}^{lat} are, of course, strictly similar. The expressions (4.10a)–(4.11b) are in agreement with those derived by Bourret.[1]

Next we will examine the three-dimensional distribution of $\gamma_{ij}(\mathbf{r}')$. First, let us consider a particular nondiagonal component, say the xy component. From (4.7) and (4.9), we obtain

$$\gamma_{xy}(\mathbf{r}')=\frac{45x'y'}{4r'^6}[3r'\coth r'+3r'^2\operatorname{csch}^2r'+2r'^3\operatorname{csch}^2r'\coth r'-8]\qquad(r'<\infty), \tag{4.12a}$$

$$\sim\frac{4}{21}x'y'+\cdots\qquad(r'<\pi). \tag{4.12b}$$

In Fig. 1, the xy section of the surface of constant $\gamma_{xy}(\mathbf{r}')$ is shown. $\gamma_{xy}(\mathbf{r}')$ is positive in the I and III quadrants and negative in II and IV quadrants, attaining peak values at four points, denoted by letters A_1, A_2, B_1, B_2 in the figure, at distances $r=\alpha r'/\pi\sim2.3\alpha/\pi$ from

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

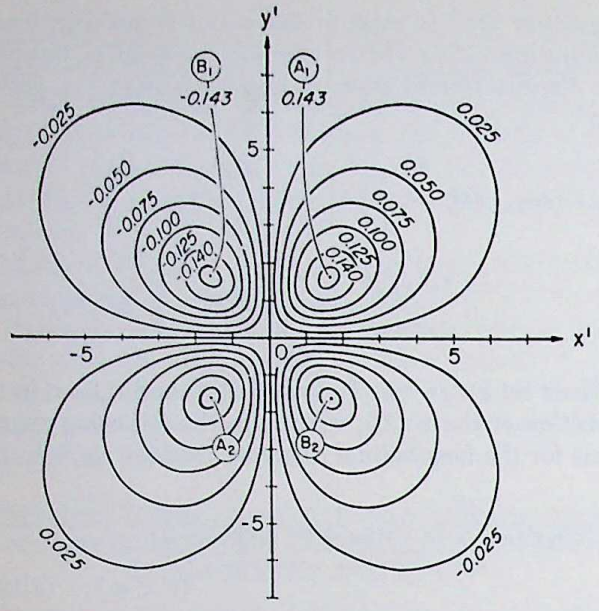


FIG. 1. Contours of $\gamma_{xy}(\mathbf{r}')$ in the xy plane. $\gamma_{xy}(\mathbf{r}')\equiv\gamma_{xy}(\mathbf{r},0)=(1/N)\,\mathcal{E}_{xy}(\mathbf{r},0)$; $N=\mathcal{E}_{xx}(0,0)=(64/45)(\pi^2/hc)^3K^4T^4$; $r=\alpha r'/\pi$; $\alpha=\hbar c/KT$.

the origin, this distance being of the order of half the mean wavelength of the radiation.[4] The variation of $\gamma_{xy}(\mathbf{r}')$ along the line $x=y$, $z=0$, is shown in Fig. 2. In Fig. 3, the contours of the surfaces $\gamma_{xy}(\mathbf{r}')$ in the plane $x=y$ are shown.

We see that the surfaces, $\gamma_{xy}(\mathbf{r}')=$constant, are closed surfaces, having the lines $x=\pm y$, $z=0$ as axes of symmetry. (Thus, any plane section through one of these lines will be symmetrical about that line.) Near the origin these surfaces are more or less hyperbolic cylinders as seen from Eq. (4.12b). The surfaces for peak values $|\gamma_{xy}(\mathbf{r}')|\simeq 0.143$, shrink to the four points A_1, A_2, B_1, B_2. The behavior of the other nondiagonal components, such as $\gamma_{yz}(\mathbf{r}')$ or $\gamma_{zx}(\mathbf{r}')$ is, of course, strictly similar.

For diagonal components of the correlation tensor, it is again sufficient to consider one of them, e.g., that

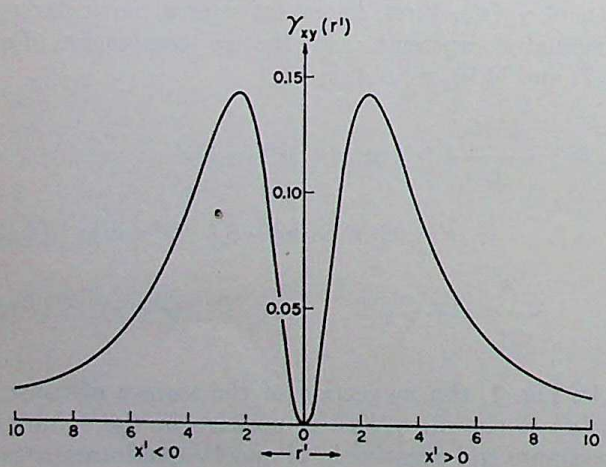


FIG. 2. Variation of $\gamma_{xy}(\mathbf{r}')$ with $\mathbf{r}'$, when $\mathbf{r}'$ is in the direction $x=y$ in the plane $z=0$.

corresponding to $i=j=z$. We have from (4.7) and (4.9)

$$\gamma_{zz}(r')=\frac{45}{4r'^4}\{(-r'\coth r'-r'^2\operatorname{csch}^2 r'-2r'^3\coth r'\operatorname{csch}^2 r'+4)+\frac{z'^2}{r'^2}(3r'\coth r'+3r'^2\operatorname{csch}^2 r'+2r'^3\coth r'\operatorname{csch}^2 r'-8)\},\quad (r'<\infty),\quad (4.13a)$$

$$=1-\frac{8}{21}r'^2+\frac{4}{21}z'^2+\cdots\qquad (r'<\pi).\quad (4.13b)$$

Eq. (4.13b) shows that for small spatial separation r' of the two points, the surfaces $\gamma_{zz}(\mathbf{r}')=$constant are the ellipsoids

$$2(x^2+y^2)+z^2=\text{constant}.$$

The contours of $\gamma_{zz}(\mathbf{r}')$ in the xy plane are circles, shown in Fig. 4. Figure 5 shows the contours in the

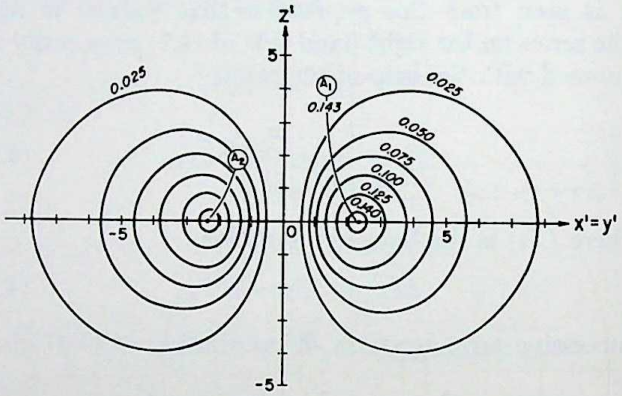


FIG. 3. Contours of $\gamma_{xy}(\mathbf{r}')$ in the plane $x=y$.

yz plane. The actual surfaces $\gamma_{zz}(\mathbf{r}')=$constant are just the surfaces of revolution generated by rotating these contours (Fig. 5) about the z axis.

For the sake of completeness the variation of γ_{zz} along the z axis (longitudinal coherence) and along the x axis (lateral coherence) are given in Figs. 6 and 7, respectively. These two curves are in agreement with Figs. 3 and 4 of Bourret.[1]

In the present section we have discussed only the case of spatial coherence characterized by $\gamma_{ij}(\mathbf{r}')\equiv\gamma_{ij}(\mathbf{r},0)$. In the general case, when τ is arbitrary, i.e., when the correlation is characterized by $\gamma_{ij}(\mathbf{r},\tau)$ rather than by $\gamma_{ij}(\mathbf{r},0)$, the value of the correlation, for any particular separation r of the two points may, of course, become appreciably larger (or smaller), provided τ is chosen appropriately.

5. THE SECOND-ORDER COMPLEX MAGNETIC AND MIXED CORRELATION TENSORS OF BLACKBODY RADIATION

Besides coherence effects associated with the electric field, coherence effects involving the magnetic field are

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

also of interest. Some of these effects are characterized by a magnetic correlation tensor $\mathcal{H}_{ij}$ and mixed correlation tensors $\mathcal{G}_{ij}$ and $\tilde{\mathcal{G}}_{ij}$, introduced in earlier publications.[16a,b] These tensors may be defined in a similar manner as the electric correlation tensor $\mathcal{E}_{ij}$. Assuming again that the field is stationary we have, in analogy with (1.3a),

$$\mathcal{H}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\langle H_i(\mathbf{r}_1,t+\tau)H_j^*(\mathbf{r}_2,t)\rangle\,, \tag{5.1}$$

$$\mathcal{G}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\langle E_i(\mathbf{r}_1,t+\tau)H_j^*(\mathbf{r}_2,t)\rangle\,, \tag{5.2}$$

$$\tilde{\mathcal{G}}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\langle H_i(\mathbf{r}_1,t+\tau)E_j^*(\mathbf{r}_2,t)\rangle\,, \tag{5.3}$$

where, as before $\mathbf{E}(\mathbf{r},t)$ denotes the complex analytic signal associated with the real electric field $\mathbf{E}^{(r)}(\mathbf{r},t)$ and $\mathbf{H}(\mathbf{r},t)$ denotes the complex analytic signal associated with the real magnetic field $\mathbf{H}^{(r)}(\mathbf{r},t)$.

We note that the tensors $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are related as follows:

$$\tilde{\mathcal{G}}_{ij}(\mathbf{r}_1,\mathbf{r}_2,\tau)=\mathcal{G}_{ji}^*(\mathbf{r}_2,\,\mathbf{r}_1,\,-\tau)\,. \tag{5.4}$$

It is known[16a,17] that *in vacuo* each of these tensors

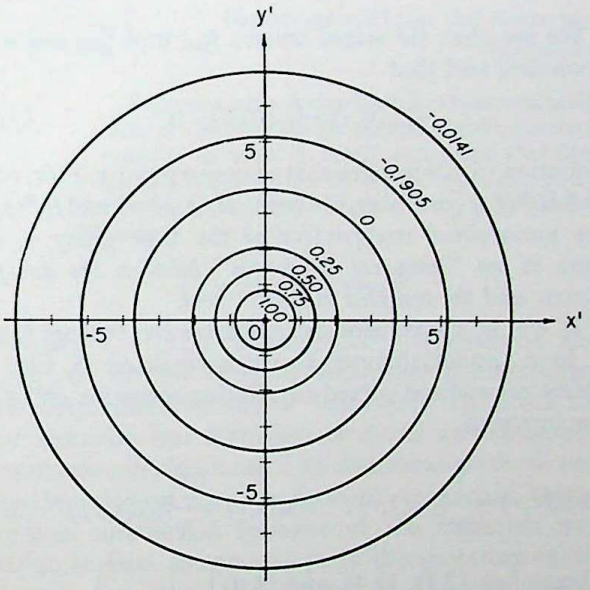


FIG. 4. Contours of $\gamma_{zz}(\mathbf{r}')$ in the xy plane.

satisfies two homogeneous wave equations. Also, that the tensors are not independent but satisfy differential equations of the form[16b,17]

$$\epsilon_{jkl}\partial_k{}^1\mathcal{E}_{lm}+\frac{1}{c}\frac{\partial}{\partial\tau}\tilde{\mathcal{G}}_{jm}=0\,, \tag{5.5}$$

$$\epsilon_{jkl}\partial_k{}^1\mathcal{G}_{lm}+\frac{1}{c}\frac{\partial}{\partial\tau}\mathcal{H}_{jm}=0\,, \tag{5.6}$$

[16] E. Wolf, (a) Nuovo Cimento **12**, 884 (1954); (b) in *Proceedings of the Symposium on Astronomical Optics*, edited by Z. Kopal (North-Holland Publishing Company, Amsterdam 1956), p. 177.
[17] P. Roman and E. Wolf, Nuovo Cimento **17**, 462 (1960).

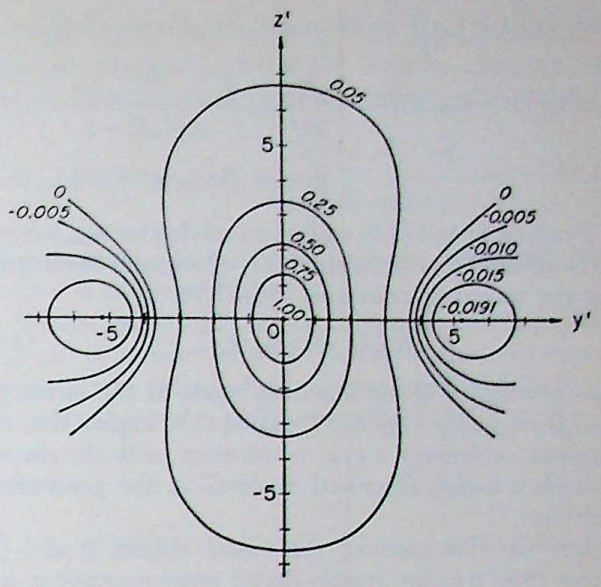


FIG. 5. Contours of $\gamma_{zz}(\mathbf{r}')$ in the yz plane. (Surfaces of constant γ_{zz} are the surfaces generated by rotation of these curves about the z axis).

etc., where $\partial_k{}^1$ $(k=x,\,y,\,z)$ are the components of the gradient, taken with respect to the coordinates of $\mathbf{r}_1$ and ϵ_{jkl} is the completely antisymmetric unit tensor of Levi-Civita. There is a similar set of differential equations involving the components $\partial_k{}^2$ of the gradient, taken with respect to the coordinates of $\mathbf{r}_2$. In the case of blackbody radiation, one has, on account of isotropy $\partial_k{}^1=-\partial_k{}^2=\partial/\partial r_k$ where r_k are components of the vector $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$.

Now for blackbody radiation we have from (2.1)–(2.4)

$$\mathcal{E}_{ij}(\mathbf{r},\tau)=\frac{hc}{2\pi^3}\int\frac{k^2\delta_{ij}-k_ik_j}{k\{\exp(\alpha k)-1\}}\times\exp\{i(\mathbf{k}\cdot\mathbf{r}-kc\tau)\}d^3k\,. \tag{5.7}$$

On substituting from (5.7) into (5.5) and solving the resulting equation for $\tilde{\mathcal{G}}$, subject to the boundary condi-

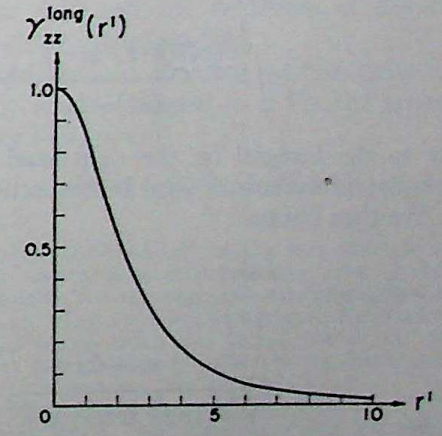


FIG. 6. Longitudinal coherence. Variation of $\gamma_{zz}{}^{\text{long}}(\mathbf{r}')\equiv\gamma_{zz}(\mathbf{r},0)$, with r', when $\mathbf{r}'$ is along the z axis.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tion $\tilde{\mathcal{G}}_{jm}(\mathbf{r},\infty)=0$, we obtain, if we also use (5.4),

$$\tilde{\mathcal{G}}_{jm}(\mathbf{r},\tau)=-\mathcal{G}_{jm}(\mathbf{r},\tau)=-\frac{hc}{2\pi^3}\epsilon_{jmk}\int\frac{k_k}{\exp(\alpha k)-1}\times\exp\{i(\mathbf{k}\cdot\mathbf{r}-kc\tau)\}d^3k\,. \quad (5.8)$$

From (5.8) and (5.6) we find on solving for $\mathfrak{C}_{jm}$ subject to the boundary condition $\mathfrak{C}_{lm}(\mathbf{r},\infty)=0$, and on comparing the resulting expression with (5.7),

$$\mathfrak{C}_{jm}(\mathbf{r},\tau)\equiv\mathcal{E}_{jm}(\mathbf{r},\tau)\,. \quad (5.9)$$

In deriving (5.9) use has been made of the identity[18] $\epsilon_{jkl}\epsilon_{mk'j}=\delta_{km}\delta_{lk'}-\delta_{kk'}\delta_{lm}$. Relation (5.9) implies that *the magnetic coherency tensor is identical with the electric coherency tensor*, discussed in detail in the preceeding sections.

Let us now consider the mixed tensors $\mathcal{G}$ and $\tilde{\mathcal{G}}$. From (5.8) we may readily derive series expansions for

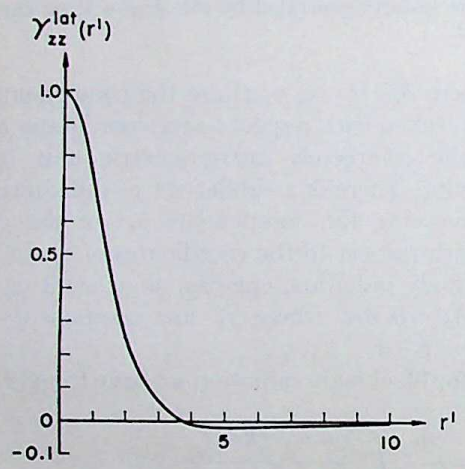


FIG. 7. Lateral coherence. Variation of $\gamma_{zz}{}^{lat}(\mathbf{r}')\equiv\gamma_{zz}(\mathbf{r},0)$, with r', when $\mathbf{r}'$ is along the x axis.

these tensors. For this purpose, we first rewrite (5.8) in the form

$$\mathcal{G}_{jm}(\mathbf{r},\tau)=-\tilde{\mathcal{G}}_{jm}(\mathbf{r},\tau)=\frac{hc}{2\pi^3}\epsilon_{jmk}(-i\partial_k)\int\frac{\exp\{i(\mathbf{k}\cdot\mathbf{r}-kc\tau)\}}{\exp(\alpha k)-1}d^3k \quad (5.10)$$

and apply to the integral on the right-hand side of (5.10) a similar procedure as used in connection with Eq. (2.5). We then obtain

$$\mathcal{G}_{jm}(\mathbf{r},\tau)=-\tilde{\mathcal{G}}_{jm}(\mathbf{r},\tau)=i\frac{16hc}{\pi^2}\epsilon_{jmk}r_k\times\sum_{n=1}^{\infty}\frac{n\alpha+ic\tau}{[(n\alpha+ic\tau)^2+r^2]^3}\,. \quad (5.11)$$

[18] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, New York, 1950), 2nd ed., p. 73.

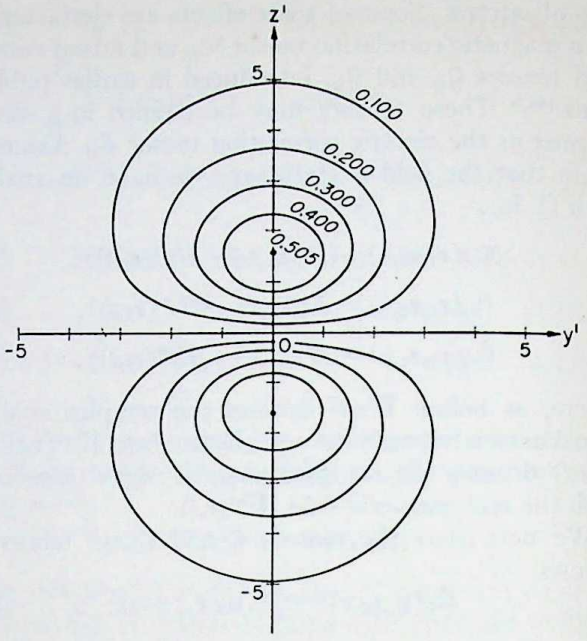


FIG. 8. Contours of $|\sigma_{xy}(\mathbf{r}')|\equiv|\sigma_{xy}(\mathbf{r},0)|$ in the yz plane. $\sigma_{xy}(\mathbf{r}')\equiv\sigma_{xy}(\mathbf{r},0)=(1/N)\mathcal{G}_{xy}(\mathbf{r},0)$; $N=(64/45)(\pi^2/hc)^3K^4T^4$.

We see that *the mixed tensors* $\mathcal{G}_{jm}$ and $\tilde{\mathcal{G}}_{jm}$ *are antisymmetric* and that

$$\mathcal{G}_{jm}(0,\tau)=\tilde{\mathcal{G}}_{jm}(0,\tau)=0\,. \quad (5.12)$$

Equation (5.12) implies that at every point $\mathbf{r}$, $E_i(\mathbf{r},t+\tau)$ and $H_j^*(\mathbf{r},t)$ [and also, of course $H_i(\mathbf{r},t+\tau)$ and $E_j^*(\mathbf{r},t)$] are *uncorrelated* irrespective of the time delay τ, i.e., *there is no "temporal coherence" between the complex electric and the complex magnetic field.*

It will be convenient to normalize the tensors $\mathcal{G}$ and $\tilde{\mathcal{G}}$ in a similar manner as we normalized $\mathcal{E}$, i.e., we define normalized mixed correlation tensors σ and $\tilde{\sigma}$ by the formulas

$$\sigma_{jm}(\mathbf{r},\tau)=\frac{1}{N}\mathcal{G}_{jm}(\mathbf{r},\tau)\,,\quad \tilde{\sigma}_{jm}(\mathbf{r},\tau)=\frac{1}{N}\tilde{\mathcal{G}}_{jm}(\mathbf{r},\tau)\,, \quad (5.13)$$

where [cf. (2.1), (2.4), and (5.9)]

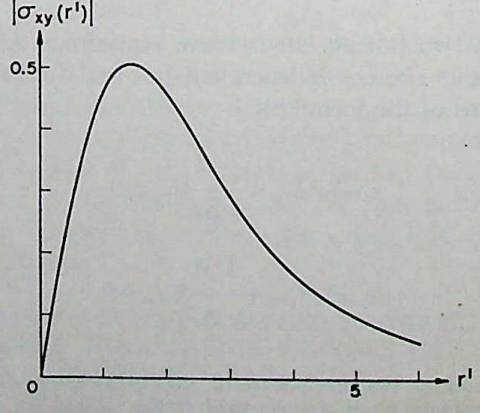


FIG. 9. Variation of $|\sigma_{xy}(\mathbf{r}')|$ with r', when $\mathbf{r}'$ is along the z axis.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$N=[\mathcal{E}_{jj}(0,0)]^{1/2}[\mathcal{H}_{jj}(0,0)]^{1/2}$$
$$=\mathcal{E}_{jj}(0,0)=\mathcal{H}_{jj}(0,0)=\frac{64\,\pi^6K^4T^4}{45\,(hc)^3}. \quad (5.14)$$

It follows from (5.13), (5.14), and (5.11) that σ and $\tilde{\sigma}$ may be represented in series form as follows:

$$\sigma_{jm}(\mathbf{r},\tau)=-\tilde{\sigma}_{jm}(\mathbf{r},\tau)=i\frac{180\alpha^4}{\pi^4}r_k\epsilon_{jmk}$$
$$\times\sum_{n=1}^{\infty}\frac{n\alpha+ic\tau}{[(n\alpha+ic\tau)^2+r^2]^3}. \quad (5.15)$$

Here the constant α is given by (2.3) as before.

From (5.15) we obtain in the special case $\tau=0$ (*spatial coherence* between the electric and magnetic fields), for typical nondiagonal elements of σ and $\tilde{\sigma}$,

$$\sigma_{xy}(\mathbf{r},0)=-\tilde{\sigma}_{xy}(\mathbf{r},0)=i\frac{180}{\alpha\pi^4}z\sum_{n=1}^{\infty}\frac{n}{(n^2+r^2/\alpha^2)^3}. \quad (5.16)$$

The diagonal elements are, of course, identically zero, since σ is antisymmetric.

We see from (5.16) that in the plane $z=0$ (xy plane) $\sigma_{xy}(\mathbf{r},0)$ is identically zero. In Fig. 8 the contours of $|\sigma_{xy}(\mathbf{r},0)|$ in the yz plane are presented and in Fig. 9 the variation of $|\sigma_{xy}(\mathbf{r},0)|$ along the z axis is shown.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Coherence Properties of Blackbody Radiation.* II. Correlation Tensors of the Quantized Field

C. L. MEHTA AND E. WOLF
Department of Physics and Astronomy, University of Rochester, Rochester, New York

(Received 9 December 1963)

Expressions are derived for the electromagnetic correlation tensors of blackbody radiation defined on the basis of the theory of the quantized field. Correlation functions of all order are considered, but second-order ones are discussed in detail; it is found that these are identical with those obtained on the basis of semiclassical theory in part I of this investigation. This result illustrates a recent theorem of E.C.G. Sudarshan relating to the equivalence between semiclassical and quantum mechanical description of statistical light beams.

1. INTRODUCTION

IN part I of this investigation,[1] expressions were derived for the complex second-order electromagnetic correlation tensors of blackbody radiation and their behavior was discussed in detail and illustrated by a number of diagrams. The statistical methods used were based entirely on classical concepts, though quantum mechanical features of the radiation were implicit in that treatment, since the spectrum of the radiation was taken to be given by Planck's law.

In the present paper the second-order correlation tensors introduced recently by Glauber[2] on the basis of the theory of the quantized field, are evaluated for blackbody radiation and are shown to be identical with those defined and evaluated on the basis of the semiclassical theory. This result illustrates a recent theorem of Sudarshan,[3] relating to the equivalence between semiclassical[4-6] and quantum mechanical description of statistical light beams.

In Sec. 4 higher-order correlation tensors of blackbody radiation are briefly discussed.

2. THE SECOND-ORDER CORRELATION TENSORS OF THE QUANTIZED FIELD

It will be useful to begin with some results which will be needed later, relating to quantization of the electromagnetic field and the definition of the correlation tensors of the quantized field.

The electric-field operator, at the space-time point $x\equiv(\mathbf{r},ct)$, when expanded in a Fourier series is given by[7] (with $i=x,\,y,\,z$)

$$\hat{E}_i(x)=\hat{E}_i^{(+)}(x)+\hat{E}_i^{(-)}(x), \quad (2.1)$$

* This research was supported by the U. S. Air Force Office of Scientific Research.

[1] C. L. Mehta and E. Wolf, Phys. Rev. 134A, 1143 (1964), preceding paper.

[2] R. J. Glauber, (a) *Electronique Quantique, 3eme Conference*, edited by N. Bloembergen and P. Grivet (Dunod Cie, Paris, 1964), p. 111; (b) Phys. Rev. 130, 2529 (1963); (c) *ibid.* 131, 2766 (1963).

[3] E. C. G. Sudarshan, (a) Phys. Rev. Letters 10, 277 (1963). (b) in *Proceedings of the Symposium on Optical Masers* (Polytechnique Press, Brooklyn, New York and John Wiley & Sons, Inc., New York, 1963), p. 45.

[4] The term "semiclassical" implies here that the distribution functions characterizing statistical properties of the beam are not necessarily non-negative and may therefore not be true probabilities. They are essentially Wigner distribution functions (see Refs. 5 and 6), called also "quasiprobabilities." However, in the present case (blackbody radiation) the distribution function turns out to be positive. [See Eq. (3.1).]

[5] E. P. Wigner, Phys. Rev. 40, 749 (1932).

[6] (a) J. E. Moyal, Proc. Cambridge Phil. Soc. 45, 99 (1949). (b) G. A. Baker, Jr., Phys. Rev. 109, 2198 (1958). (c) C. L. Mehta, J. Math. Phys. 5, 677 (1964).

[7] All operators are denoted by circumflex.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$\hat{E}_i^{(+)}(x)=i\left(\frac{2\pi\hbar c}{V}\right)^{1/2}\sum_{\mathbf{k}}(k)^{1/2}\{\sum_{\lambda=1}^{2}\epsilon_i^{(\lambda)}\hat{a}_\lambda(\mathbf{k})e^{i\kappa x}\}, \quad (2.2)$$

$$\hat{E}_i^{(-)}(x)=\{\hat{E}_i^{(+)}(x)\}^\dagger=-i\left(\frac{2\pi\hbar c}{V}\right)^{1/2}\sum_{\mathbf{k}}(k)^{1/2}$$
$$\times\{\sum_{\lambda=1}^{2}\epsilon_i^{(\lambda)}\hat{a}_\lambda^\dagger(\mathbf{k})e^{-i\kappa x}\}. \quad (2.3)$$

Here V represents the volume to which the field is confined and $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}_\lambda^\dagger(\mathbf{k})$ are the annihilation and creation operators, respectively, for a photon of momentum $\hbar\mathbf{k}$ and polarization λ; they satisfy the commutation relations

$$[\hat{a}_\lambda(\mathbf{k}),\hat{a}_{\lambda'}^\dagger(\mathbf{k}')]=\delta_{\lambda\lambda'}\delta_{\mathbf{k}\mathbf{k}'}, \quad (2.4)$$

$$[\hat{a}_\lambda(\mathbf{k}),\hat{a}_{\lambda'}(\mathbf{k}')]=[\hat{a}_\lambda^\dagger(\mathbf{k}),\hat{a}_{\lambda'}^\dagger(\mathbf{k}')]=0. \quad (2.5)$$

$\boldsymbol{\varepsilon}^{(1)}$ and $\boldsymbol{\varepsilon}^{(2)}$ are two unit vectors, such that $\boldsymbol{\varepsilon}^{(1)}$, $\boldsymbol{\varepsilon}^{(2)}$ and $\mathbf{k}/k$ (with $k=|\mathbf{k}|$) form a right-handed orthogonal triad and κx represents the usual four-vector product

$$\kappa x=\mathbf{k}\cdot\mathbf{r}-kct. \quad (2.6)$$

In writing down Eqs. (2.2) and (2.3), the modified Lorentz condition was used in the form introduced by Gupta and Bleuler,[8] employing an indefinite metric: Allowable states $|\psi\rangle$ are only those for which

$$\sum_{\mu=1}^{4}\frac{\partial\hat{A}_\mu^{(+)}}{\partial x_\mu}|\psi\rangle=0, \quad (2.7)$$

where $\hat{A}^{(+)}$ is the positive frequency part of the potential of the field. This condition justifies our restriction to transverse photons ($\lambda=1,2$) as long as only expectation values of products of field operators are considered.

The magnetic-field operator may be similarly written in the form

$$\hat{H}_i(x)=\hat{H}_i^{(+)}(x)+\hat{H}_i^{(-)}(x), \quad (2.8)$$

where

$$\hat{H}_i^{(+)}(x)=i\left(\frac{2\pi\hbar c}{V}\right)^{1/2}\sum_{\mathbf{k}}(k)^{1/2}$$
$$\times\left\{\sum_{\lambda=1}^{2}\frac{(\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)})_i}{k}\hat{a}_\lambda(\mathbf{k})e^{i\kappa x}\right\}, \quad (2.9)$$

$$\hat{H}_i^{(-)}(x)=\{\hat{H}_i^{(+)}(x)\}^\dagger=-i\left(\frac{2\pi\hbar c}{V}\right)^{1/2}\sum_{\mathbf{k}}(k)^{1/2}$$
$$\times\left\{\sum_{\lambda=1}^{2}\frac{(\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)})_i}{k}\hat{a}_\lambda^\dagger(\mathbf{k})e^{-i\kappa x}\right\}. \quad (2.10)$$

If $\hat{\rho}$ is the density operator of the field, the second-order correlation tensor of the electric field, introduced by Glauber[2a,b] may defined as

$$'\mathcal{E}_{ij}(x_1,x_2)=\mathrm{Tr}\{\hat{\rho}\hat{E}_i^{(-)}(x_1)\hat{E}_j^{(+)}(x_2)\}, \quad (2.11)$$

where the indices i, j specify Cartesian components of the tensor (i, $j=x$, y, z) and the prime on $\mathcal{E}$ is used to distinguish this tensor from the corresponding one defined in paper I in classical terms. In a similar way, one may define the second-order magnetic and mixed-correlation tensors

$$'\mathcal{H}_{ij}(x_1,x_2)=\mathrm{Tr}\{\hat{\rho}\hat{H}_i^{(-)}(x_1)\hat{H}_j^{(+)}(x_2)\}, \quad (2.12)$$

$$'\mathcal{G}_{ij}(x_1,x_2)=\mathrm{Tr}\{\hat{\rho}\hat{E}_i^{(-)}(x_1)\hat{H}_j^{(+)}(x_2)\}, \quad (2.13)$$

$$'\tilde{\mathcal{G}}_{ij}(x_1,x_2)='\mathcal{G}_{ji}^*(x_2,x_1)$$
$$=\mathrm{Tr}\{\hat{\rho}\hat{H}_i^{(-)}(x_1)\hat{E}_j^{(+)}(x_2)\}. \quad (2.14)$$

To evaluate the correlation tensors $'\mathcal{E}$, $'\mathcal{H}$, $'\mathcal{G}$, and $'\tilde{\mathcal{G}}$ for blackbody radiation, it will be useful first of all to express the density operator $\hat{\rho}$ in a representation in which the base vectors are the eigenvectors $|\{z\}\rangle$ of the annihilation operator $\hat{a}_\lambda(\mathbf{k})$:

$$\hat{a}_\lambda(\mathbf{k})|\{z\}\rangle=z_\lambda(\mathbf{k})|\{z\}\rangle. \quad (2.15)$$

Here $\{z\}$ denotes the set of $z_\lambda(\mathbf{k})$ for all λ's and $\mathbf{k}$'s, so that

$$|\{z\}\rangle=\prod_{\lambda,\mathbf{k}}|z_\lambda(\mathbf{k})\rangle. \quad (2.16)$$

Since the operator $\hat{a}_\lambda(\mathbf{k})$ is not Hermitian, its eigenvalues $z_\lambda(\mathbf{k})$ are in general complex,

$$z_\lambda(\mathbf{k})=x_\lambda(\mathbf{k})+iy_\lambda(\mathbf{k}), \quad (2.17)$$

($x_\lambda(\mathbf{k})$, $y_\lambda(\mathbf{k})$ real) and the eigenvectors belonging to different eigenvalues are not orthogonal. However, they obey a closure relation (cf. Ref. 9).

$$\frac{1}{\pi}\int\prod_{\lambda,\mathbf{k}}|z_\lambda(\mathbf{k})\rangle\langle z_\lambda(\mathbf{k})|\,d^2z_\lambda(\mathbf{k})=\hat{1}, \quad (2.18)$$

where $\hat{1}$ is the identity operator if, as we shall assume, they are normalized so that

$$\langle z_\lambda(\mathbf{k})|z_\lambda(\mathbf{k})\rangle=1. \quad (2.19)$$

The eigenstate $|z_\lambda(\mathbf{k})\rangle$ has the following expansion in the number representation[9]:

$$|z_\lambda(\mathbf{k})\rangle=\sum_{n_\lambda(\mathbf{k})=0}^{\infty}\exp(-\tfrac{1}{2}|z_\lambda(\mathbf{k})|^2)\frac{\{z_\lambda(\mathbf{k})\}^{n_\lambda(\mathbf{k})}}{[n_\lambda(\mathbf{k})!]^{1/2}}|n_\lambda(\mathbf{k})\rangle, \quad (2.20)$$

where the $|n_\lambda(\mathbf{k})\rangle$ form an orthonormal set of eigenvectors of the number operator $\hat{a}_\lambda^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k})$:

$$\hat{a}_\lambda^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k})|n_\lambda(\mathbf{k})\rangle=n_\lambda(\mathbf{k})|n_\lambda(\mathbf{k})\rangle. \quad (2.21)$$

The eigenvectors of the annihilation operator have been found very useful in the analysis of problems

[8] See, for example, S. S. Schweber, *An Introduction to Quantum Field Theory* (Harper & Row, New York, 1961), p. 242.

[9] J. R. Klauder, Ann. Phys. (N. Y.) 11, 123 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

relating to the statistical behavior of radiation.[2,3,10] An important property of these eigenvectors is expressed by a theorem established recently by Sudarshan[3] which we shall employ. According to this theorem, any operator and in particular the density operator $\hat{\rho}$ may be expressed in the "diagonal" form

$$\hat{\rho}=\int \phi(\{z\})\,|\{z\}\rangle\langle\{z\}|\,d^2\{z\}\,. \tag{2.22}$$

Using this representation of the density operator, it then follows from (2.2), (2.3), (2.11), (2.22) and (2.15) and its Hermitian adjoint that the second-order correlation tensor of the electric field may be expressed in the form

$$'\mathcal{E}_{ij}(x_1,x_2)$$

$$=\int \phi(\{z\})\langle\{z\}|\hat{E}_i^{(-)}(x_1)\hat{E}_j^{(+)}(x_2)|\{z\}\rangle d^2\{z\}$$

$$=\sum_{\lambda,\mathbf{k}}\sum_{\lambda',\mathbf{k}'}\epsilon_i^{(\lambda)}(\mathbf{k})\epsilon_j^{(\lambda')}(\mathbf{k}')\Phi_{\lambda\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)\,, \tag{2.23}$$

where

$$\Phi_{\lambda\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)=(2\pi\hbar c/V)(kk')^{1/2}\exp(-i\kappa x_1+i\kappa' x_2)$$

$$\times\int \phi(\{z\})z_\lambda^*(\mathbf{k})z_{\lambda'}(\mathbf{k}')d^2\{z\}\,. \tag{2.24}$$

The corresponding expressions for the second order magnetic and mixed correlation tensors defined by (2.12)–(2.14) can also readily be written down. We only need to use the expressions (2.9) and (2.10) in place of (2.2) and (2.3) where appropriate. We then obtain in place of (2.23) the following expressions for the other second-order correlation tensors of the quantized field:

$$'\mathcal{H}_{ij}(x_1,x_2)=\sum_{\lambda,\mathbf{k}}\sum_{\lambda',\mathbf{k}'}\left(\frac{\mathbf{k}\times\boldsymbol{\epsilon}^{(\lambda)}}{k}\right)_i\left(\frac{\mathbf{k}'\times\boldsymbol{\epsilon}^{(\lambda')}}{k'}\right)_j$$

$$\times\Phi_{\lambda\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)\,, \tag{2.25}$$

$$'\mathcal{G}_{ij}(x_1,x_2)=\sum_{\lambda,\mathbf{k}}\sum_{\lambda',\mathbf{k}'}\epsilon_i^{(\lambda)}(\mathbf{k})\left(\frac{\mathbf{k}'\times\boldsymbol{\epsilon}^{(\lambda')}}{k'}\right)_j$$

$$\times\Phi_{\lambda\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)\,, \tag{2.26}$$

$$'\tilde{\mathcal{G}}_{ij}(x_1,x_2)=\sum_{\lambda,\mathbf{k}}\sum_{\lambda',\mathbf{k}'}\left(\frac{\mathbf{k}\times\boldsymbol{\epsilon}^{(\lambda)}}{k}\right)_i\epsilon_j^{(\lambda')}(\mathbf{k}')$$

$$\times\Phi_{\lambda\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)\,. \tag{2.27}$$

3. THE SECOND-ORDER CORRELATION TENSORS OF BLACKBODY RADIATION

We will now evaluate the correlation tensors $'\mathcal{E}$, $'\mathcal{H}$, $'\mathcal{G}$ and $'\tilde{\mathcal{G}}$ for the case of blackbody radiation. It would appear from a discussion of Glauber [Ref. 2c, Eq. (10.23)], based essentially on the central limit theorem, that the "phase-space" distribution function $\phi(\{z\})$ for blackbody radiation is given by

$$\phi(\{z\})=\prod_{\lambda,\mathbf{k}}\frac{1}{\pi\bar{n}_k}\exp\left\{-\frac{|z_\lambda(k)|^2}{\bar{n}_k}\right\}, \tag{3.1}$$

where $\bar{n}_k\equiv\bar{n}_{\mathbf{k},\lambda}$ $(\lambda=1,2)$ is the average number of photons with momentum $\mathbf{k}$ and polarization λ,

$$\bar{n}_k=1/(e^{\alpha k}-1)\,, \tag{3.2}$$

with

$$\alpha=\hbar c/KT\,. \tag{3.3}$$

That the "phase-space" distribution $\phi(\{z\})$ given by (3.1) does indeed lead to the well-known expression for the density operator $\hat{\rho}$ of a radiation field in thermal equilibrium at temperature T is verified directly in the Appendix to the present paper (see also Mandel[10]).

If we substitute from (3.1) into (2.24), and also use the relation

$$\int z_{\lambda'}^*(\mathbf{k}')z_{\lambda''}(\mathbf{k}'')\prod_{\lambda,\mathbf{k}}\frac{1}{\pi\bar{n}_k}\exp(-|z_\lambda(\mathbf{k})|^2/\bar{n}_k)d^2z_\lambda(\mathbf{k})$$

$$=\frac{1}{\pi\bar{n}_{k'}}\delta_{\lambda'\lambda''}\delta_{\mathbf{k}'\mathbf{k}''}\int|z_{\lambda'}(\mathbf{k}')|^2$$

$$\times\exp(-|z_{\lambda'}(\mathbf{k})|^2/\bar{n}_{k'})d^2z_{\lambda'}(\mathbf{k}')$$

$$=\bar{n}_{k'}\delta_{\lambda'\lambda''}\delta_{\mathbf{k}'\mathbf{k}''}\,, \tag{3.4}$$

where δ is the Kronecker symbol, we find that

$$\Phi_{\lambda,\lambda'}(\mathbf{k},\mathbf{k}';x_1,x_2)$$

$$=\frac{2\pi\hbar c}{V}k\bar{n}_k\delta_{\lambda\lambda'}\delta_{\mathbf{k}\mathbf{k}'}\exp\{-i\kappa(x_1-x_2)\}\,. \tag{3.5}$$

Next we assume that the linear dimensions of the enclosure are large compared with the mean wavelength of the radiation. Summation over $\mathbf{k}$ may then be replaced by integration over the whole $\mathbf{k}$ space, if use is made of the usual rules[11]

$$\frac{1}{(V)^{1/2}}\sum_{\mathbf{k}}f(\mathbf{k})\rightarrow\frac{1}{(2\pi)^{3/2}}\int d^3k f(\mathbf{k})\,, \tag{3.6}$$

$$\delta_{\mathbf{k}\mathbf{k}'}\rightarrow\delta(\mathbf{k}-\mathbf{k}')\,, \tag{3.7}$$

where f is an arbitrary function and $\delta(\mathbf{k}-\mathbf{k}')$ is the three-dimensional Dirac delta function. One then obtains from (3.5) and (2.23), the following expression for the correlation tensor $'\mathcal{E}$:

$$'\mathcal{E}_{ij}(x_1,x_2)=\frac{\hbar c}{4\pi^2}\int k\bar{n}_k\exp\{i\kappa(x_2-x_1)\}$$

$$\times\sum_{\lambda=1}^{2}\epsilon_i^{(\lambda)}(\mathbf{k})\epsilon_j^{(\lambda)}(\mathbf{k})d^3k\,. \tag{3.8}$$

[10] L. Mandel, Phys. Letters 7, 117 (1963).

[11] J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison Wesley Publishing Co., Cambridge, 1955), p. 38.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Next we substitute for $\bar{n}_k$ from (3.2) and also use the relation

$$\sum_{\lambda=1}^{2} \epsilon_i^{(\lambda)}(\mathbf{k})\epsilon_j^{(\lambda)}(\mathbf{k}) = \delta_{ij} - \frac{k_i k_j}{k^2}, \tag{3.9}$$

which follows from the fact that $\boldsymbol{\varepsilon}^{(1)}$, $\boldsymbol{\varepsilon}^{(2)}$, and $\mathbf{k}/k$ form a triad of mutually orthogonal unit vectors. We thus finally obtain the following expression for the second-order electric correlation tensor $'\mathcal{E}$ of the quantized field:

$$'\mathcal{E}_{ij}(x_1,x_2) = \frac{\hbar c}{4\pi^2}\int \frac{k^2\delta_{ij} - k_i k_j}{k(e^{ak}-1)} \times \exp\{i\kappa(x_2-x_1)\}d^3k. \tag{3.10}$$

For the magnetic second-order correlation tensor, we obtain in a strictly similar manner from (2.25)

$$'\mathcal{H}_{ij}(x_1,x_2) = \frac{\hbar c}{4\pi^2}\int k\bar{n}_k \exp\{i\kappa(x_2-x_1)\} \times \sum_{\lambda=1}^{2}\left(\frac{\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)}}{k}\right)_i\left(\frac{\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)}}{k}\right)_j d^3k. \tag{3.11}$$

Now the three vectors $\mathbf{k}\times\boldsymbol{\varepsilon}^{(1)}/k$, $\mathbf{k}\times\boldsymbol{\varepsilon}^{(2)}/k$ and $\mathbf{k}/k$ also form a triad of mutually orthogonal unit vectors, and hence a relation similar to (3.9) holds between them. In consequence, the right-hand side of (3.11) becomes identical with the right-hand side of (3.10) and hence we have

$$'\mathcal{H}_{ij}(x_1,x_2) = '\mathcal{E}_{ij}(x_1,x_2). \tag{3.12}$$

For the mixed second-order correlation tensor $'\mathcal{G}$ we obtain from (2.26) in a similar manner

$$'\mathcal{G}_{ij}(x_1,x_2) = \frac{\hbar c}{4\pi^2}\int k\bar{n}_k \exp\{i\kappa(x_2-x_1)\} \times \sum_{\lambda=1}^{2}\epsilon_i^{(\lambda)}(\mathbf{k})\left(\frac{\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)}}{k}\right)_j d^3k. \tag{3.13}$$

It may readily be verified that

$$\sum_{\lambda=1}^{2}\epsilon_i^{(\lambda)}(\mathbf{k})\left(\frac{\mathbf{k}\times\boldsymbol{\varepsilon}^{(\lambda)}}{k}\right)_j = \epsilon_{ijl}k_l/k, \tag{3.14}$$

where ϵ_{ijl} is the completely antisymmetric unit tensor of Levi Civita. Hence, if also (2.14) is used, (3.13) reduces to

$$'\mathcal{G}_{ij}(x_1,x_2) = -'\tilde{\mathcal{G}}_{ij}(x_1,x_2) = \frac{\hbar c}{4\pi^2}\epsilon_{ijl}\int \frac{k_l}{e^{ak}-1} \times \exp\{i\kappa(x_2-x_1)\}d^3k. \tag{3.15}$$

Noting that $\kappa(x_2-x_1)$ is the four-vector product

$$\kappa(x_2-x_1) = kc\tau - \mathbf{k}\cdot\mathbf{r}, \tag{3.16}$$

where $\tau = t_1 - t_2$ and $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, it follows on comparing Eqs. (3.10), (3.12), and (3.15) with Eqs. (5.7), (5.9), and (5.8) of paper I that

$$\begin{aligned} '\mathcal{E}_{ij}(x_1,x_2) &= \tfrac{1}{4}\mathcal{E}_{ij}^*(x_1,x_2), \\ '\mathcal{H}_{ij}(x_1,x_2) &= \tfrac{1}{4}\mathcal{H}_{ij}^*(x_1,x_2), \\ '\mathcal{G}_{ij}(x_1,x_2) &= \tfrac{1}{4}\mathcal{G}_{ij}^*(x_1,x_2), \\ '\tilde{\mathcal{G}}_{ij}(x_1,x_2) &= \tfrac{1}{4}\tilde{\mathcal{G}}_{ij}^*(x_1,x_2). \end{aligned} \tag{3.17}$$

The relations (3.17) show that the second-order electromagnetic correlation tensors of blackbody radiation defined for the quantized field by Eqs. (2.11), (2.12), (2.13), and (2.14) are proportional to the complex conjugates of the corresponding tensors defined for the classical field by Eqs. (1.3a), (5.1), (5.2), and (5.3) of paper I.[12–14]

4. HIGHER ORDER CORRELATION TENSORS

Up to now we have considered correlation tensors of second order only. For the sake of completeness, we will now briefly consider electromagnetic correlation tensors of higher orders for the case of blackbody radiation. Again these correlation tensors may be defined either as appropriate averages of products involving the field vectors of the complex classical field at a number of space-time points,[15,16,17] or as quantum mechanical expectation values involving the corresponding field operators.[2] Since the equivalence of these different definitions has been demonstrated by Sudarshan,[3] Mandel[10] and Mehta and Wolf,[18] we may restrict our discussion to correlation tensors defined on the basis of the theory of the quantized field. The electric correlation tensor of order $m+n$ is then defined by the equation

$$\begin{aligned} &\mathcal{E}^{(m,n)}_{j_1,j_2,\cdots j_{m+n}}(x_1,x_2,\cdots x_m;\, x_{m+1},\cdots x_{m+n}) \\ &\quad = \mathrm{Tr}\{\hat{\rho}\hat{E}_{j_1}^{(-)}(x_1)\hat{E}_{j_2}^{(-)}(x_2)\cdots\hat{E}_{j_m}^{(-)}(x_m) \\ &\qquad \times \hat{E}_{j_{m+1}}^{(+)}(x_{m+1})\cdots\hat{E}_{j_{m+n}}^{(+)}(x_{m+n})\}, \end{aligned} \tag{4.1}$$

where $\hat{E}_j^{(-)}(x)$, $\hat{E}_j^{(+)}(x)$ are Cartesian components of the operators $\hat{E}^{(-)}(x)$, $\hat{E}^{(+)}(x)$ at the space-time point

---

[12] The correlation tensors of the classical field could readily be redefined without any loss of generality, so as to lead to strict identities $'\varepsilon_{ij} = \varepsilon_{ij}$, etc. In particular, the factor $\frac{1}{4}$ could be suppressed by employing a slightly different normalization in defining the analytic signal (Refs. 13 and 14) by means of which the complex fields are associated with the real fields. However, we preferred to retain the customary definitions throughout this investigation.

[13] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, London and New York, 1959), Chap. X.

[14] D. Gabor, J. Inst. Elec. Engrs. **93**, Part III, 429 (1946).

[15] L. Mandel, in *Electronique Quantique 3eme Conference*, edited by N. Bloembergen and P. Grivet (Dunod Cie, Paris, 1964), p. 101.

[16] E. Wolf in *Electronique Quantique 3 eme Conference*, edited by N. Bloembergen and P. Grivet (Dunod Cie, Paris, 1964), p. 13.

[17] E. Wolf, in *Proceedings of the Symposium on Optical Masers* (Polytechnique Press, Brooklyn, New York and John Wiley & Sons, Inc., New York, 1963), p. 29.

[18] C. L. Mehta and E. Wolf (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

x, defined by Eqs. (2.2) and (2.3). It follows from Sudarshan's discussion[3] that the expectation value on the right of (4.1) may be expressed as a phase-space average of the products of the components of the complex classical field, in the form

$$\mathcal{G}^{(m,n)}_{j_1,j_2,\cdots j_{m+1}}(x_1,x_2,\cdots x_m;\, x_{m+1},\cdots x_{m+n})$$

$$=\int \phi(\{z\})E_{j_1}{}^*(x_1)E_{j_2}{}^*(x_2)\cdots E_{j_m}{}^*(x_m)$$

$$\times E_{j_{m+1}}(x_{m+1})\cdots E_{j_{m+n}}(x_{m+n})d^2\{z\}\,, \quad (4.2)$$

where $\phi(\{z\})$ is the "phase space" distribution function associated with the density operator $\hat{\rho}$ [cf. (2.22)] and

$$E_j(x)=i\left(\frac{2\pi\hbar c}{V}\right)^{1/2}\sum_{\mathbf{k},\lambda}(k)^{1/2}\epsilon_j{}^{(\lambda)}(\mathbf{k})z_\lambda(\mathbf{k})e^{ikx}; \quad (4.3)$$

$(j=x,y,z)$ is the eigenvalue of the operator $\hat{E}_j{}^{(+)}(x)$ corresponding to the eigenstate $|\{z\}\rangle$.

In the case of blackbody radiation, the phase-space distribution $\phi(\{z\})$ is given by (3.1), which is a multivariate Gaussian distribution. Now since according to (4.3) the field components $E_{j_\mu}(x_\mu)$, $E_{j_\mu}{}^*(x_\mu)$ are linear combinations of the $z_\lambda(\mathbf{k})$ and $z_\lambda{}^*(\mathbf{k})$, they will, according to a well-known theorem on the Gaussian random process, be also distributed as Gaussian variates[19] and hence all moments involving them may completely be expressed in terms of the second-order moments[20] as follows:

If $m=n$

$$\mathcal{G}^{(m,n)}_{j_1,j_2,\cdots j_{m+n}}(x_1,x_2,\cdots x_m;\, x_{m+1},\cdots x_{m+n})$$

$$=\sum\nolimits_\pi \mathcal{G}^{(1,1)}_{j_1,j_p}(x_1;\,x_p)\,\mathcal{G}^{(1,1)}(x_2;\,x_q)\cdots$$

$$\times\,\mathcal{G}^{(1,1)}_{j_m,j_s}(x_m;\,x_s)\,. \quad (4.4)$$

If $m\neq n$

$$\mathcal{G}^{(m,n)}_{j_1,j_2,\cdots j_{m+n}}(x_1,x_2,\cdots x_m;\, x_{m+1},\cdots x_{m+n})=0. \quad (4.5)$$

The second-order coherence tensor $\mathcal{G}^{(1,1)}_{j_1,j_p}(x_1;\,x_p)$, etc. on the right of (4.4) is precisely the second order electric correlation tensors given by (3.10) and π denotes all permutations p, q, $\cdots s$ of the non-negative integers $m+1$, $m+2$, $\cdots 2m$.

Strictly similar expressions can readily be written down for the magnetic and mixed tensors of an arbitrary order.

Finally, it should be mentioned that the expression (4.4) for the electric correlation tensor of blackbody radiation has also recently been derived by Glauber.[20a]

[19] Proof of this result for a real Gaussian random process is given, for example, in Wang and Uhlenbeck, Rev. Mod. Phys. 17, 323 (1945). Generalization to a complex Gaussian random process is straightforward.

[20] See, for example, I. S. Reed, Inst. Radio Engrs. Trans. Inform. Theory IT8, 194 (1962).

APPENDIX

In this Appendix we will verify that the phase-space distribution (3.1), namely,

$$\phi(\{z\})=\prod_{\lambda,\mathbf{k}}\frac{1}{\pi\bar{n}_k}\exp\left\{-\frac{|z_\lambda(\mathbf{k})|^2}{\bar{n}_k}\right\} \quad (A1)$$

corresponds to the density operator $\hat{\rho}$ for radiation field in thermal equilibrium at temperature T.

On substituting from (A1) and (2.20) into (2.22), we obtain

$$\hat{\rho}=\prod_{\lambda,\mathbf{k}}\sum_{n_\lambda(\mathbf{k})=0}^{\infty}\sum_{m_\lambda(\mathbf{k})=0}^{\infty}\int\frac{1}{\pi\bar{n}_k}$$

$$\times\exp\left\{-\frac{|z|^2}{\bar{n}_k}\right\}\frac{z^{n_\lambda(\mathbf{k})}z^{*m_\lambda(\mathbf{k})}}{(n_\lambda(\mathbf{k})!m_\lambda(\mathbf{k})!)^{1/2}}$$

$$\times\exp\{-|z|^2\}|n_\lambda(\mathbf{k})\rangle\langle m_\lambda(\mathbf{k})|\,d^2z. \quad (A2)$$

Next we set $z=r\exp(i\theta)$, $d^2z=rdrd\theta$ and note that the integration over θ gives $2\pi\delta_{n_\lambda(\mathbf{k}),m_\lambda(\mathbf{k})}$. We then obtain

$$\hat{\rho}=\prod_{\lambda,\mathbf{k}}\sum_{n_\lambda(\mathbf{k})}\int_0^\infty\frac{1}{\bar{n}_k}\exp\left\{-r^2\left(1+\frac{1}{\bar{n}_k}\right)\right\}$$

$$\times(r^2)^{n_\lambda(\mathbf{k})}\frac{|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})|}{n_\lambda(\mathbf{k})!}d(r^2)$$

$$=\prod_{\lambda,\mathbf{k}}\sum_{n_\lambda(\mathbf{k})}\frac{1}{\bar{n}_k+1}\left(1+\frac{1}{\bar{n}_k}\right)^{-n_\lambda(\mathbf{k})}|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})|\,. \quad (A3)$$

Next we substitute for $\bar{n}_k$ from (3.2) and find that

$$\hat{\rho}=\prod_{\lambda,\mathbf{k}}\frac{\exp(\alpha k-1)}{\exp(\alpha k)}\sum_{n_\lambda(\mathbf{k})}\exp\{-n_\lambda(\mathbf{k})\alpha k\}$$

$$\times|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})|\,, \quad (A4)$$

where α is given by (3.3). But

$$\sum_{n_\lambda(\mathbf{k})}\exp\{-n_\lambda(\mathbf{k})\alpha k\}|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})|$$

$$=\exp\{-\hat{a}_\lambda{}^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k})\alpha k\}\sum_{n_\lambda(\mathbf{k})}|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})| \quad (A5)$$

and we have also the completeness relation

$$\sum_{n_\lambda(\mathbf{k})}|n_\lambda(\mathbf{k})\rangle\langle n_\lambda(\mathbf{k})|=\hat{1}. \quad (A6)$$

Using (A5) and (A6) in (A4), we finally obtain the following expression for the density operator $\hat{\rho}$ corresponding to the phase-space distribution function (A1):

$$\hat{\rho}=\prod_{\lambda,\mathbf{k}}\frac{\exp\{-\alpha k\hat{a}_\lambda{}^\dagger(\mathbf{k})\hat{a}_\lambda(\mathbf{k})\}}{\exp(\alpha k)/[\exp(\alpha k)-1]}. \quad (A7)$$

This is precisely the density operator for a radiation field in thermal equilibrium at temperature T.[21]

[21] A. Messiah, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961), Vol. I, p. 448.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Effect of Magnetic Field on Thermal Conductivity and Energy Gap of Superconducting Films*

D. E. Morris† and M. Tinkham
Department of Physics, University of California, Berkeley, California

(Received 9 January 1964)

We have measured the change in the thermal conductivity of superconducting tin, indium, and lead films upon application of a magnetic field in the plane of the film. These experiments were undertaken to explore the dependence of the energy gap upon magnetic field and to determine the thermodynamic order of the field-induced transition in films. Within the range of temperatures available to us ($0.35\ T_c$ to $0.65\ T_c$) the thermal conductivity of indium and tin films increases nearly as H^2. In films of thickness $d \geqq 2800$ Å the thermal conductivity jumps at H_c to the normal state value, indicating a first-order phase transition in thin films, consistent with the Ginzburg-Landau (GL) theory. If the effect of a field upon the superconducting state can be adequately represented as a change of the ϵ_0 of BCS, we may use the theory of Bardeen, Rickayzen, and Tewordt to compute $\epsilon_0(H)$ from our data. At $T=0.65\ T_c$ we find $\epsilon_0(H)/\epsilon_0(H=0)=[1-(H/H_c)^2]^{1/2}$ in agreement with the GL prediction. At $T=0.35\ T_c$ however, $\epsilon_0(H)/\epsilon_0(0)=1-(H/H_c)^2$ provides a more satisfactory fit to our data. Orientation of the field $\perp$ and $\parallel$ to the direction of heat flow produces the same effect on the thermal conductivity to $\pm 1\%$. This is interesting in view of Bogoliubov's prediction of a $\mathbf{p}_F \cdot \mathbf{v}_{\text{drift}}$ term in the excitation spectrum of a current-carrying superconductor. When the field is not parallel to the surface of the film the thermodynamic transition is still sharply defined, but H_c is reduced. Thin lead films gave results similar to those of indium and tin films except that at low fields the field-dependent conductivity increased more slowly than H^2 at low temperatures.

I. INTRODUCTION

IT is now well accepted on both experimental[1,2] and theoretical[3,4] grounds that near T_c the superconducing transition of a thin film in a magnetic field is of second order. The predictions of the Ginzburg-Landau-Gor'kov theory[3–6] (GLG), applicable here, are in good agreement with experimental results. At reduced temperatures $t=T/T_c \ll 1$, however, the GLG theory is inapplicable.[7] Bardeen[8] has made a calculation which leads to a first-order transition at very low temperatures, while Nambu and Tuan[9] find a second-order transition for a very thin film at all temperatures.

The elegant experimental measurements of the field dependence of the energy gap employing the tunnel effect[1,2] have generally been limited to a few metals at moderately high reduced temperatures and to the case of fields applied parallel to the film surface. Only recently, while this report was in preparation, have measurements been extended to lead films at low reduced temperatures by Douglass and Meservy.[10] The seexperiments support our findings, to be reported here, that the transition for a thin film remains second order even at very low temperatures, and they also find a similar field dependence of the energy gap at low temperatures. It should perhaps be noted that the tunnel-effect measurements selectively indicate the effective energy gap at the film surface for electrons moving normal to the surface.

We have developed an independent experimental method to study the thermodynamic parameters of a superconducting film in a magnetic field. This method consists of measuring the effect of an externally applied magnetic field on the thermal conductivity of a superconducting thin film. A preliminary report of our results on tin and indium films has been published elsewhere.[11] The experiments were then extended[12] to a wider range of experimental conditions and to another metal, lead. We report these more extensive experiments here. The experimental data reveal the thermodynamic order of the transition by inspection. By using the relation derived by Bardeen, Rickayzen, and Tewordt[13] (BRT) the data may be inverted point by point to give $\epsilon_0(H)$, the field-dependent energy gap averaged through the film.

The external field has been applied at various angles with respect to the plane of the film with unexpected results, and the relative directions of the field and the heat flow in the plane of the film have been varied in a search for vestiges of the anisotropy of the energy gap

* Supported in part by the A. P. Sloan Foundation, the National Science Foundation, and the Office of Naval Research.

† Present address: Morris Electronics (P) Ltd., Poona, India.

[1] I. Giaever and K. Megerle, Phys. Rev. 122, 1101 (1961).

[2] D. H. Douglass, Jr., Phys. Rev. Letters 7, 14 (1961).

[3] V. L. Ginzburg and L. D. Landau, Zh. Eksperim. i Teor. Fiz. 20, 1064 (1950).

[4] D. H. Douglass, Jr., Phys. Rev. Letters 6, 346 (1961).

[5] L. P. Gor'kov, Zh. Eksperim. i Teor. Fiz. 36, 1918 (1959) [English transl.: Soviet Phys.—JETP 9, 1364 (1959)].

[6] V. L. Ginzburg, Zh. Eksperim. i Teor. Fiz. 34, 113 (1958) [English transl.: Soviet Phys.—JETP 7, 78 (1958)].

[7] V. L. Ginzburg, Zh. Eksperim i Teor. Fiz. 30, 593 (1956) [English transl.: Soviet Phys.—JETP 3, 621 (1956)].

[8] J. Bardeen, Rev. Mod. Phys. 34, 667 (1962).

[9] Y. Nambu and S. F. Tuan, Phys. Rev. Letters 11, 119 (1963); Phys. Rev. 133, A1 (1964).

[10] D. H. Douglass, Jr., and L. M. Falicov, *Progress in Low Temperature Physics*, edited by C. J. Gorter (North-Holland Publishing Company, Amsterdam, to be published).

[11] D. E. Morris and M. Tinkham, Phys. Rev. Letters 6, 600 (1961).

[12] D. E. Morris, Ph.D. thesis, University of California, Berkeley, 1962 (unpublished).

[13] J. Bardeen, G. Rickayzen, and L. Tewordt, Phys. Rev. 113, 982 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

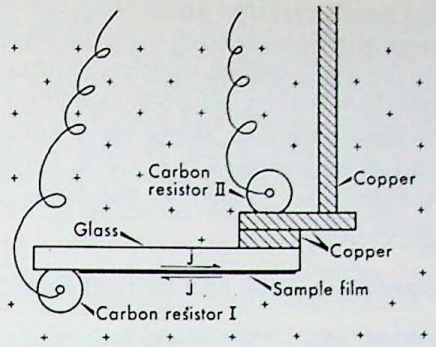


FIG. 1. Sample and thermometer mounting arrangement.

in a current-carrying superconductor predicted by Bogoliubov.[14]

II. EXPERIMENTAL METHOD

A. Low-Temperature Arrangement

The use of carbon resistance thermometers,[15] with their low thermal inertia and high sensitivity, makes it possible to measure the low-temperature thermal conductivity of thin metal films deposited on glass substrates. During our experiments, the temperature drop along the film is typically 0.1°K with 10^{-7} W total heat transport along the sample film and substrate. Under these conditions a change of 2 μdeg in the temperature drop along the film is detectable.

The sample and thermometer arrangement is shown in Fig. 1. A $\frac{1}{2}$-W carbon-composition resistor is prepared for use as a thermometer-heater (I) by grinding a flat surface until the resistance material just becomes visible. The lead wires are cut off short, and lengths of No. 42 Formvar insulated manganin wire are attached with nonsuperconducting solder. The thermal conductivity of these lead wires is negligible in our experiments. The flattened surface of the resistor is glued with GE-7031 cement at one end of the glass sample-film substrate. Using the same cement, the opposite end of the substrate is attached to a copper heat sink, which is in turn connected to the lid of the experimental chamber. A similarly prepared second resistor, to be used as a reference thermometer (II), is similarly glued to the copper heat sink.

The experimental chamber is placed in the lowest part of the interior of a conventional, liquid-nitrogen shielded, glass Dewar for liquid helium use. A thin-walled stainless-steel tube extends from the experimental chamber to the top of the Dewar. This tube is used to evacuate the chamber; it also carries No. 40 manganin wires for connection to the thermometers. Since the usual power input to the heater-thermometer during the experiments is only $\sim 10^{-7}$ W, thermal anchoring of the manganin leads to the inside of the chamber lid and radiation baffling in the pumping tube above the chamber are both essential. The chamber is sealed to its copper lid with a Woods metal joint. When this alloy is cooled below ~8°K, it becomes superconducting.[16] In order to prevent the superconducting diamagnetism and trapped flux at the joint from disturbing the magnetic field applied to the sample film, the sample is suspended near the bottom of the $4\frac{1}{2}$-in.-long chamber from a copper rod directed downward from the copper lid.

B. Thermometry

The thermometers used were Allen-Bradley $\frac{1}{2}$-W nominal 100-Ω resistors, connected in a wheatstone-bridge circuit. Satisfactory cancellation of temperature drifts was achieved without special matching of the reference and heater resistors. The resistance of the reference thermometer was measured as a function of the vapor pressure of the helium bath over the range 1.2–4.2°K at a small enough current to prevent appreciable heating above the bath temperature. The results could be approximated closely by a relation of the form $\ln R = A + B/T$, where $A = 2.89$ and $B = 3.26°$.

The insensitivity of the resistance thermometers to magnetic fields was checked during several experimental runs. If the magnetic field is increased to several times the critical field of the experimental film, any changes in the thermometer bridge balance above the critical field may be attributed to magnetoresistance in the thermometers. This is so because any effect of the magnetic field on the thermal resistance of the experimental films, which have electron mean free paths ~800 Å, would in the normal state be negligible even at the maximum field applied. In the case of a film of relatively low critical field the resistance bridge balance was unaffected by an increase from $H = 400$ Oe to $H = 800$ Oe. In another experiment, increase of the field from 7500 to 15 000 Oe produced only a small shift in the resistance thermometer bridge balance. These results are consistent with the behavior of similar resistors as determined by Clement and Quinnell.[15] They found that magnetoresistance effects were small and proportional to H^2. Further, we expect the small magnetoresistance of the thermometer heater and the reference thermometer to nearly cancel in measurements of resistance bridge balance.

The relation between power input and resistance change was measured separately for the heater thermometer and the reference thermometer in several experiments. The temperature rise of the reference thermometer at a given power input was only about 10% of that of the heater thermometer at the same power level. We may conclude that the thermal resistance of the re-

[14] N. N. Bogoliubov, Zh. Eksperim. i Teor. Fiz. 34, 58 (1958) [English transl.: Soviet Phys.—JETP 7, 41 (1958)].
[15] J. R. Clement and E. H. Quinnell, Rev. Sci. Instr. 23, 213 (1952).
[16] D. Shoenberg, *Superconductivity* (Cambridge University Press, Cambridge, 1952).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sistors themselves and of the glue used to fasten them in place amounts to $\sim\frac{1}{10}$ of the thermal resistance of the sample film and substrate. For example, at 1.2° and a power input of 0.09×10^{-6} W, the temperature rise of the heater thermometer was ~0.060° and of the reference ~0.006°. This gives a value of $dP/dT=1.5\times10^{-5}$ W/deg for the reference thermometer. Berman[17] established that in the 100-Ω ½-W Allen-Bradley resistor (the same type and value as we have used) $dP/dT=5.2\times10^{-5}$ W/deg. The difference is accounted for by the fact that thermal contact was made to only a fraction of the surface of the resistor in our experiments.

C. Temperature Stability

The initial (unsuccessful) experiments were attempted at 1.3°K, without a reference thermometer. A monotonic, but not perfectly uniform, decrease of the bath temperature at the rate of ~1 mdeg/min during the measurements resulted from the decreasing helium film creep as the liquid-helium level declined. The resulting drift in the thermometer resistance overwhelmed the temperature changes expected from the desired effect of a magnetic field. Attempts were made to stabilize the bath temperature with an electronic regulator consisting of a separate resistance thermometer in a 33-cps ac bridge, followed by a tuned amplifier, lock-in detector, and dc heater in the liquid helium. Although satisfactory long-term stability was achieved, amplifier noise in the heater produced objectionable variations over periods shorter than a minute. Since the total experimental effect is ~5 mdeg, the noise level should be less than 50 μdeg to ensure an adequate signal/noise ratio.

With the introduction of the reference thermometer in a bridge circuit with the thermometer heater, the effective drift rate was decreased by a factor of 10, even without special matching of the resistors. This permitted data to be taken without the use of any noise-producing feedback stabilization schemes.

Temperatures of 1.2°K were reached by pumping at 300 liters/min with a conventional rotary vacuum pump. Operation at 1.7°K was attained by adjustment of a valve in the pumping line until the temperature drift rate at the desired temperature was reasonably low. When operation near 2.1°K was desired, a Cartesian manostat was connected in the helium vapor pumping line. A 1-cu-ft ballast tank was connected in the pumping line between the Dewar and the manostat to reduce fluctuations in the pressure above the helium bath which were caused by the oscillation of the manostat throttle.

Spatial nonuniformity of the bath temperature and bubbling of the helium usually results in intolerable temperature fluctuations of the sample if operation is attempted above the λ point of the helium bath. However, it was possible to conduct experiments at the normal boiling point (4.22°K) by leaving the bath completely undisturbed and isolated from vibration for three hours after the liquid-helium transfer. By that time, the bubbling produced only a relatively small rapid fluctuation of the signal; however, the temperature drop along the film had to be increased to ~0.6°K (several times the low-temperature values) to obtain an adequate signal-to-noise ratio.

D. Temperature Drop and Critical Field Spread

The approximate temperature drop along the film during the experiments at various temperatures is shown in Table I. The increase in ΔT at higher temperatures was necessary to maintain the signal/noise ratio in the presence of greater uncontrolled temperature fluctuations and decreasing thermometer sensitivity. The percentage difference in the critical fields at the hot and cold ends of the film which result from the difference of temperature is also shown in Table I. These values were determined from the appropriate bulk critical field curves.

TABLE I. Temperature drop along experimental films and corresponding spread in critical fields.

Mean temperature	ΔT	$\Delta H_c/H_c$ Indium	Tin	Lead
1.2°	0.06°	1.5%	1.0%	0.2%
1.7°	0.1°	4%	3.5%	0.6%
2.1°	0.15°	10%	7%	1.3%
4.5°	0.6°			15%

E. Magnetic Field

A high-power water-cooled Helmholtz pair produced fields up to 1000 Oe. This relatively high-field, iron-free system was employed to permit precise control of the direction of the field by using auxiliary trimming coils, in order that the effect of applying the field slightly out of the plane of the sample film could be investigated. The terminals of the trimming coils could be connected by a rotary switch to taps on a low-resistance nichrome strip wired in series with the main coil pair. In this way the angle of the resultant field, independent of amplitude, could be adjusted in 0.02° steps up to ±4°. In experiments on films with critical fields greater than 1000 Oe, conventional iron-core electromagnets were used. The resulting uncontrolled angle of the field with respect to the plane of the film is less than 1° and apparently it has no major effect upon the experimental results for these very thin films, as may be seen in Fig. 11.

F. Sample Films

Sample films were produced by rapid evaporation onto glass substrates held at 77°K, after thorough ion-bombardment cleaning of the substrate surface.

[17] R. Berman, Rev. Sci. Instr. 25, 94 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The substrates were No. 0 microscope-cover glasses about 0.11 mm thick which had been broken into pieces about 5 mm×15 mm, suitable for our experiments. The substrates were washed in a saturated solution of detergent, rinsed in hot tap water, and air dried. The substrate was then clamped against the surface of an unfilled cold trap inside the vacuum chamber. The chamber was evacuated and an intense (20 mA) high-voltage (5000 V) glow discharge was operated for 5 min to heat and clean the substrate by ion bombardment. Then the pressure was reduced to 3×10^{-6} mm Hg and the substrate was cooled to 77°K by filling the attached cold trap. The vacuum pumping line contained a liquid nitrogen cooled baffle to rule out contamination of the vacuum chamber by condensation of oil vapor.

A tantalum evaporation boat containing the preweighed and pre-fused sample charge was rapidly heated to about 1400°C and the film evaporation took place in a period ranging from 1 sec for a 3-mg charge (giving a 1300-Å film) to 6 sec for a 30-mg charge. The source-to-substrate distance was only 2 in. The deposition rate of over 1000 Å/sec, along with the residual gas pressure of less than 10^{-5} mm Hg, ensured that the flux of residual gas molecules was a great deal smaller than that of the metal atoms during the evaporation. Films prepared under these conditions should therefore be quite pure. Cooling of the substrate during the evaporation is necessary to prevent migration of the metal atoms after they first strike the surface.[18] After the films were warmed to room temperature and removed from the vacuum, they were annealed at room temperature for at least one week. (The lead films were annealed in vacuum to retard oxidation.)

The sources and purities of the metals used in evaporation of the experimental films are given in Table II.

TABLE II. Metal sources and purities.

Metal	Supplier	Stated purity
Indium	Indium Corporation of America	99.999%
Lead	A. D. Mackay	99.999%
Tin	Johnson, Matthey	99.999%

Films ranging in thickness from 250 to 12 000 Å were prepared. Thickness was estimated from the measured critical field and published data relating the critical field to film thickness. As described above, all films were deposited on substrates cooled to 77°K. In all cases it was possible to prepare films which were mirror-like in appearance with very little diffuse reflectivity, in contrast to results of room-temperature evaporations.[18] No structure was visible under magnification of 1500 diameters except a fine granularity.

A 250-Å indium film showed no noticeable deterioration after seven months exposure to the atmosphere. A 700-Å tin film similarly exposed for four months showed no visible deterioration. Oxidation of lead films, however, was more rapid. The experimental film designated Pb I was exposed to the atmosphere for approximately 5 h while Pb II was exposed approximately ½ h. In order to determine the durability of lead films and the extent of deterioration produced by atmospheric exposure while preparing for an experiment, measurements have been made of the transmission of 0.9-μ radiation through a lead film of about 180-Å thickness, as a function of the duration of exposure to the atmosphere. The results may be roughly interpreted as a decrease in thickness of the metallic film due to oxidation, resulting in decreased reflection and increased transmission. We assume that Z, the impedance per square, is inversely proportional to d, the film thickness. Then, for transmission $T\ll1$, $T\sim Z^2\sim1/d^2$. On this basis, during the period between 15-min and 2 h after the beginning of exposure to the atmosphere, Δd was ~7 Å. After one day Δd was ~20 Å, and the oxidation of the film proceeded at ~7 Å per day for the next week. Now, PbO is a semiconductor, not an insulator as the above analysis has assumed; also, the estimation of d_0 from the weight of the evaporated Pb charge may be somewhat in error. Nevertheless we may safely conclude that the atmospheric oxidation of the Pb films used in our experiments was of negligible importance.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A. Electron Mean Free Path

An estimate of the electron mean free path Λ_e in the experimental films will be useful in the interpretation of our data. According to BRT[13] the mean free path of the electronic excitations is unaffected by the transition to superconductivity. We may determine Λ_e in the normal state from the relation for thermal conductivity $K_n=\frac{1}{3}C_{en}v_F\Lambda_e$ where the electronic specific heat $C_{en}\approx4\times10^{-4}\times T$ cal/mole-deg for both tin and indium,[19] and $v_F\approx10^8$ cm/sec is the Fermi velocity. A direct measurement of K_n is impractical because of the dominant thermal conductance of the glass substrate. It is possible, however, to determine experimentally the ratio of the thermal conductivity of the films to that of the substrate and combine this information with published values of the thermal conductivity of glass, K_{gl}, in order to find K_n. We will denote by r the fractional change in the total conductance of the film plus substrate combination when the film is driven into the normal state by a field $H\geq H_c$. The quantity r is measured experimentally as the change in the thermometer bridge balance position upon application of $H\geq H_c$ (i.e., the total "signal") divided by the shift in bridge balance position resulting from increase in the bridge power in-

[18] L. Holland, *Vacuum Deposition of Thin Films* (John Wiley & Sons, Inc., New York, 1956), pp. 206–208.

[19] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc., New York, 1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

put from a small value up to the full operating current. When r is combined with the published values[20] of the temperature dependent ratio K_s/K_n of the thermal conductivities of the superconducting and normal states,[16] and with d and D, the thickness of the film and of the substrate, we may determine $K_n/K_{gl}=rD/(1-K_s/K_n)d$. Then $\Lambda_e=3K_n/C_{en}v_F$. The thermal conductivity of the glass, K_{gl}, needed in this calculation, is estimated from a value[19] of $K_{gl}=5\times10^{-4}$ W/cm-deg at 2°K, and the assumption that $K_{gl}\sim T^{1.3}$ at liquid-helium temperatures.[21] The resulting values for K_{gl} were used in the computation of the Λ_e values quoted in Table III.

TABLE III. Data leading to estimates of electronic mean free paths Λ_e in the experimental films. Symbols are defined in text. In all cases thickness of glass substrate $D=0.11$ mm.

Sample	d (Å)	T (°K)	Heater power (μW)	r	$\frac{K_s}{K_n}$	K_{gl} (mW/cm-deg)	K_{gl}' (mW/cm-deg)	Λ_e (Å)
SnII	2600	1.2°	0.09	0.2	0.1	0.26	0.11	600
SnIII	700	1.2°	0.08	0.046	0.1	0.26	0.07	500
InII	1600	1.7°	0.3	0.07	0.39	0.40	0.18	600
		1.8°	0.1	0.07	0.45	0.40	0.20	600
		2.1°	0.22	0.11	0.61	0.55	0.73	1600
		2.14°	0.85	0.07	0.64	0.58	0.5	800
InIII	650	1.25°	0.14	0.05	0.16	0.26	0.2	600
		1.7°	0.11	0.07	0.33	0.40	0.12	1000
		2.1°	1.3	0.016	0.65	0.55	0.58	600
InIV	4000	1.2°	0.09	0.41	0.14	0.26	0.26	800
		2.1°	1.3	0.10	0.67	0.55	0.44	600

[a] From Ref. 16, p. 83, Fig. 29.

Note that Λ_e has a value of 600–900 Å, and seems to be constant with increasing film thickness between 650 Å (InIII) and 4000 Å (InIV). This indicates that Λ_e is limited by imperfection and impurity scattering rather than boundary scattering in these films. In either case, Λ_e would be independent of temperature. Thus, the range of calculated values of Λ_e at different temperatures in the same film gives an indication of the lack of precision in the determination of Λ_e.

Phonon scattering of the electrons is negligible in our films. The phonon-limited K_{eg} would be $2.56\times10^3/T^2$ W/cm-deg in tin, according to the compilation of Klemens.[22] From $K_e=(1/3)\gamma T v_F\Lambda_e$ we may calculate at 1.2°K, $\Lambda_{eg}\approx0.4$ cm, and at 2.1°K, $\Lambda_{eg}\approx0.075$ cm.

An independent experimental estimate of the thermal conductivity of the glass substrate, K_{gl}', also listed in Table III, was made from the power input to the heater, the resulting temperature change in the thermometer resistance, and the approximate dimensions of the substrate. We ascribe the difference between K_{gl} and K_{gl}' at the lower temperatures to poor calibration of the thermometer resistance as a function of temperature. This calibration does not affect the normalization or interpretation of the experimental data; it was used only in the computation of K_{gl}' and in the estimation of the temperature difference between the ends of the film during the experiments (Table I).

B. Temperature Variation of H_c/H_{cb}

We note an increase in H_c/H_{cb} with temperature in all of the experimental films, (see Figs. 2, 5–8). This may be anticipated as a result of the increase in λ with temperature, since H_c/H_{cb} is determined by the ratio λ/d. The effect has been studied in detail by Appleyard *et al.*[23] and by Alekseyevsky[24]; their results are discussed by Shoenberg.[16] Our measurements of H_c are limited to temperature $T<0.7\ T_c$ where the changes of λ are relatively small, and our results are only of sufficient accuracy to show an agreement in sign and order of magnitude with the expected effect. Our results fall between $(H_c-H_{cb})/H_{cb}\sim\lambda(T)\sim[1-(T/T_c)^4]^{-1/2}$, which applies for thick films ($d\gg\lambda$), and $H_c/H_{cb}(T)\sim[1-(T/T_c)^2]^{1/2}$, predicted by Pippard[25] for very small specimens.

Pippard[25] and Ginzburg[6] have predicted substantial superheating and supercooling effects in films a few penetration depths thick. In several experiments we approached the transition from above and also from below and found no detectable hysteresis. This may be a result of the thermal gradient along our films, or of their crystalline imperfection.

C. Behavior of Very Thin Films

The results of our experiments are easiest to interpret in the case of very thin films ($d<\lambda$). The applied field penetrates almost uniformly, and ϵ_0 is also expected to be essentially uniform.

Figure 2 shows results for thin (~700 Å) films of tin and indium. For both metals, the thermal conductivity K increases nearly as H^2 up to the critical field, the thermal conductivity in the superconducting state joins smoothly onto the field-independent conductivity of the normal metal. The location of the shoulder at this point fixes H_c for the pruposes of normalizing the data. The absence of any discontinuity at H_c indicates that the superconducting transition in a magnetic field is second order for films of this thickness, in contrast to the first-order transition of bulk superconductors in a field. In other words, since K is related

[20] For tin and indium, K_s/K_n as a function of T/T_c, measured in bulk samples, is in good agreement with the theoretical calculation of BRT for the case of electronic thermal conductivity with lattice-defect scattering dominant. Boundary scattering would give essentially similar results. In the films we have studied, even if we neglect defect scattering, the electronic mean free path would be limited by boundary scattering to a few thousand angstroms at most, and thus we can have considerable confidence in the applicability of the experimental and theoretical bulk metal values of K_s/K_n.

[21] D. Bijl, Physica 14, 684 (1949).

[22] P. G. Klemens, *Encyclopedia of Physics* (Springer-Verlag, Berlin, 1956), Vol. XIV, p. 245.

[23] E. T. S. Appleyard, J. R. Bristow, H. London, and A. D. Misener, Proc. Roy. Soc. (London) A172, 540 (1939).

[24] N. E. Alekseyevsky, J. Phys. USSR 4, 401 (1941).

[25] A. B. Pippard, Phil. Mag. 43, 273 (1952).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

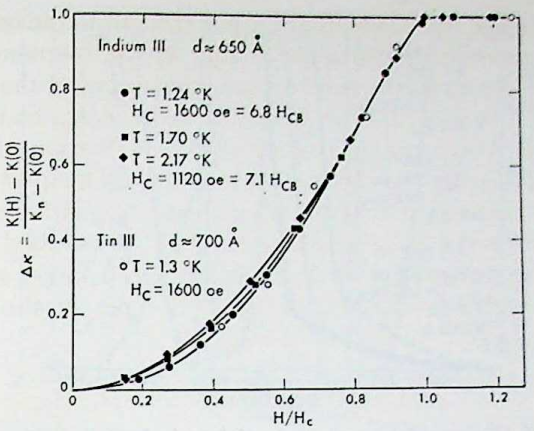


FIG. 2. Change of thermal conductivity of thin superconducting films with field. Note that the phase transition is of second order for films of this thickness.

directly to the gap width, we can infer that the gap closes continuously to zero when a field is applied to a thin film, whereas in a bulk sample it drops discontinuously from nearly the full value (~98% even at the surface) to zero.

The measured dependence of K upon H may be used to compute a dependence of the energy gap upon field in a simple way if the electronic term K_e is primarily limited by scattering of electrons by lattice imperfections. Under these conditions the ratio of the thermal conductivity in the superconducting state K_{es} to that in the normal state K_{en} is given by Bardeen, Rickayzen, and Tewordt[13] as their Eq. (3.6). This may be written

$$\frac{K_{es}}{K_{en}} = G(\epsilon_0/kT) = \int_{\epsilon_0}^{\infty} E^2(\partial f/\partial E)dE \Big/ \int_0^{\infty} E^2(\partial f/\partial E)dE,$$

where $f(E/kT)$ is the Fermi function. The function $G(\epsilon_0/kT)$ is plotted in Fig. 3. If we assume that the effect of a field upon the superconducting state can be adequately represented as a change of the ϵ_0 of BCS, then the observed quantity $\Delta\kappa$ plotted in Fig. 2 should be given by

$$\Delta\kappa = \frac{K_{es}(H)-K_{es}(0)}{K_{en}-K_{es}(0)} = \frac{G[\epsilon_0(H)/kT]-G[\epsilon_0(0)/kT]}{1-G[\epsilon_0(0)/kT]}.$$

This relation was used to invert the experimental data on indium III, point-by-point, to yield $\epsilon_0(H)$, with the result shown in Fig. 4. Because $G(\epsilon_0/kT)$ initially drops only cubically as ϵ_0/kT increases from zero (see Fig. 3), the gap must drop more steeply to zero near H_c than the approach of $K_{es}(H)$ to K_{en} there. A rather steep final drop $\sim(H_c-H)^{1/2}$ in the gap near H_c is necessary to reproduce the almost linear approach of K_{es} to K_{en} observed at that point.

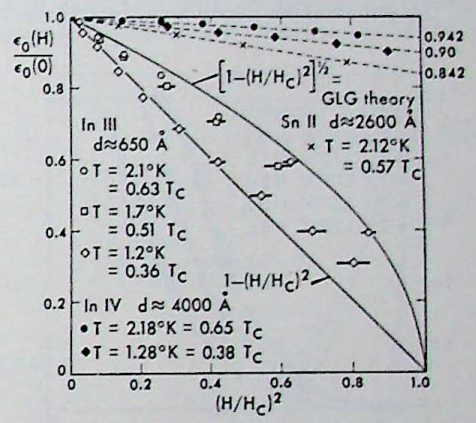


FIG. 4. Magnetic field dependence of superconducting energy gap for several tin and indium films computed from the data in Figs. 2, 6, 7, and 8, assuming that the energy gap is independent of position in the film. The curve $[1-(H/H_c)^2]^{1/2}$ is that given by the GLG theory in the limit of very thin films.

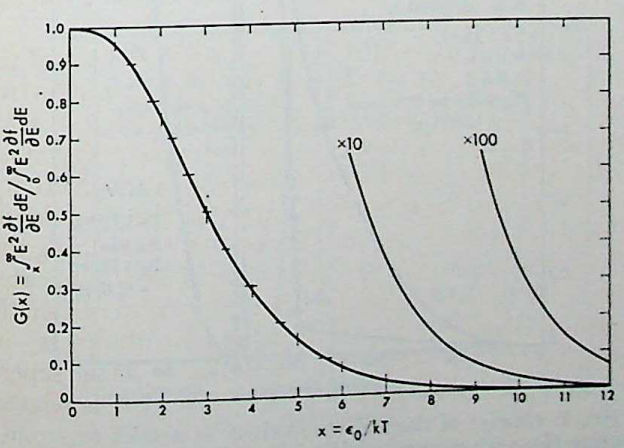


FIG. 3. The function $G(\epsilon_0/kT)$ relates the normalized thermal conductivity to the energy gap, according to the theory of Bardeen, Rickayzen, and Tewordt.

The difficulty of distinguishing between a broadened transition because of film inhomogeneity and a truly rounded approach to K_{en} because of the form of $G(\epsilon_0/kT)$ results in an uncertainty in the choice of H_c. The bars on the plotted experimental points in Fig. 4 indicate the size of this uncertainty.

Our results for the dependence of the energy gap upon magnetic field in a thin film are qualitatively similar to those from the more direct tunnel experiments of Giaever and Megerle[1] and of Douglass,[2] although their data were taken at higher reduced temperature and in a different metal (aluminum). From our data shown in Fig. 4, we see that $\epsilon_0(H)/\epsilon_0(0)$ seems to approach $1-h^2$, where $h=H/H_c$, at low temperatures, but it is moderately well fitted at $T/T_c=0.63$ by $(1-h^2)^{1/2}$. The latter form is that given by Douglass[4] based on the Ginzburg-Landau-Gor'kov (GLG) theory, which is expected to hold near T_c. We also note that our data at $T/T_c=0.36$ deviate from $1-h^2$ and approach $(1-h^2)^{1/2}$ when the gap has dropped so that $\epsilon_0(H)<kT$. This behavior indicates that the GLG approach becomes successful, as expected,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

when there is a large amount of thermal excitation present.

We may find the approximate value of d/λ for In III from the GLG theory. Ginzburg[6] has shown that the solution of the Ginzburg-Landau equations, in the limit $d<\lambda$, leads to $H_c=(24)^{1/2}(\lambda/d)H_{cb}$. For In III, $H_c/H_{cb}\approx 7$, so that $d\approx 0.7\lambda$. The sevenfold increase in critical field over the bulk value shows that the field penetration must be very nearly complete, so that we are in this case observing the effect of a field in the thin film limit.

D. Behavior of Thicker Films

The film indium II (Fig. 5) of intermediate thickness $d\approx 1800$ Å, is still thin enough to display a second-order transition, but the increase of K near H_c is much faster than H^2. This behavior may be explained in terms of an increase of the penetration depth λ as the gap decreases,

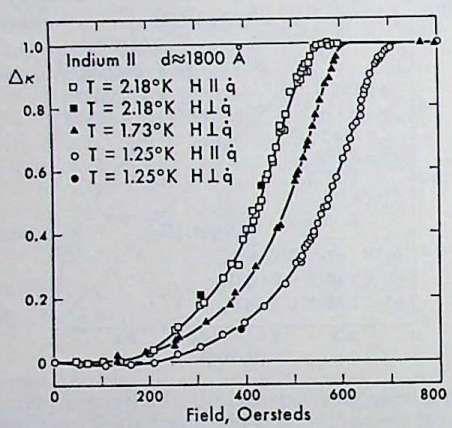


FIG. 5. Change of thermal conductivity of a superconducting film of intermediate thickness with magnetic field.

as would be expected on the basis of sum-rule arguments.[26] Since the film thickness is comparable to the electron mean free path (≈ 600 Å in this sample according to Table III) which determines the coherence length ξ, we expect the energy gap to be only weakly dependent on position in the film. We may then apply the approach of Pippard,[25] considering the balance between condensation energy and magnetic energy to determine the gap (or order parameter) as a function of field. Taking into account the dependence of film susceptibility on (d/λ) and the dependence of λ on ϵ_0, one is led to expect a change in the form of $\epsilon_0(H/H_c)$ of the observed sense for films with $d>\lambda$. A similar conclusion follows from the GLG theory, as discussed by Douglass.[4]

In Fig. 6 we see the experimental results on a moderately thick film, tin II. The data taken at 2.12° shows a variation of K with H which is purely quadratic (within 1%) up to the critical field. At this point $\Delta\kappa=24\%$. With a further increase of field, K increases rapidly

[26] M. Tinkham and R. A. Ferrell, Phys. Rev. Letters 2, 331 (1959).

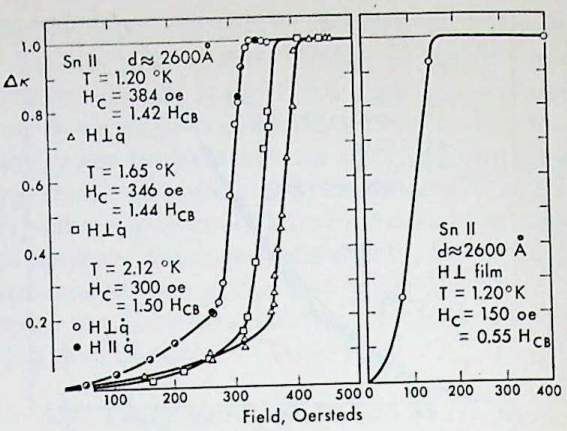


FIG. 6. Change of thermal conductivity of a 2600-Å tin film with magnetic field at various temperatures. Because of poor temperature stabilization, we have little confidence in the 1.20 and 1.65°K curves except for the determination of critical fields. The right-hand curve is for H perpendicular to the film plane.

up to the normal-state conductivity. The breadth of the transition can be accounted for partly, but not completely, by the difference in temperature of the ends of the film. We do not have much confidence in the 1.2 and 1.65° curves, except for the points which determine the transition, because the temperature stabilization was poor during those particular runs.

When we apply our method for calculation of the energy gap as a function of field to the 2.12° data, assuming a uniform gap throughout the film, we find the energy gap decreasing exactly as H^2 to $\epsilon_0(H)/\epsilon_0(0)=0.842$ at the critical field (Fig. 4). If we apply the GLG theory we find that the observed critical field ratio $H_c/H_{cb}=1.50$ implies that $d/\lambda=3.6$ and $\epsilon_0(H_c)/\epsilon_0(0)=0.78$. The agreement with the above value of 0.842 is quite good considering that at H_c the gap in the interior of the film should be depressed somewhat less than that at the surface, because in this film the coherence length is limited by Λ_e to a value less than the film thickness.

We next consider the data for In IV (Figs. 7, 8, 9, 10).

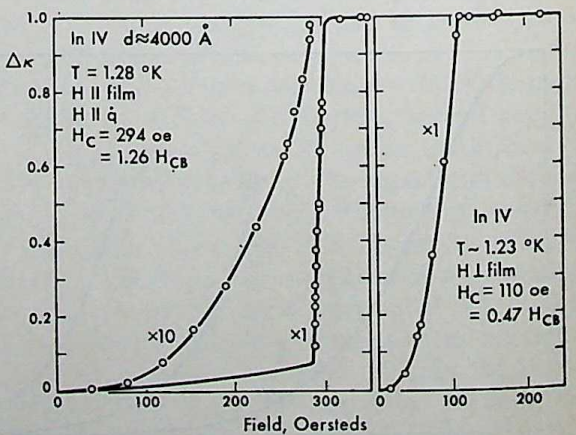


FIG. 7. Change of thermal conductivity of a thick superconducting film with magnetic field, at $T\approx 0.38\ T_c$. The first points are shown on an expanded scale for greater clarity. At the right is the result of applying the field perpendicular to the surface of the film.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

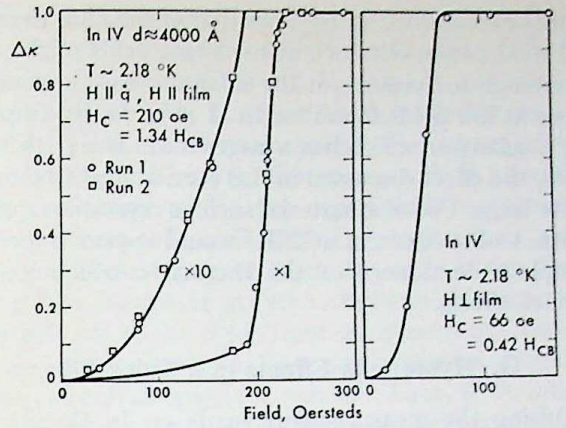


FIG. 8. Change of thermal conductivity of a thick superconducting film with magnetic field, at $T \approx 0.64\ T_c$.

Because of the large film thickness the change in thermal conductivity does not exceed about 10% up to the critical field, and the calculated energy gap variation assuming ϵ_0 is independent of position, i.e., $\xi \gg d$ (Fig. 4) is correspondingly small. The numbers at the right-hand margin of Fig. 4 are found by extrapolating the computed values of the energy gap to H_c at the points shown. These points fall almost exactly on straight lines until we reach the transition region, so we may be confident of the extrapolation. There seems to be a very slight curvature to the $0.38\ T_c$ curve, probably reflecting a trace of the peculiar hysteresis effects which are so prominent in this sample with a field orientation **H** perpendicular to $\dot{\mathbf{q}}$, the direction of heat flow (Figs. 9, 10).

E. Initial Decrease of $\Delta\kappa$

In indium II (Fig. 5) at the lowest temperature we see an apparent initial decrease of thermal conductivity with field. This may be a spurious effect arising from a slow approach to the equilibrium distribution of any trapped flux. We note from Table III that the heater power in our experiments is only about 10^{-6}–10^{-7} W and that r, the total change in thermal conductance of the metal and glass-sample sandwich with magnetic field is between 5 and 40%. Thus, evolution of heat at a rate of only 10^{-10} W (<4 erg/h!) when the field is applied could produce a spurious effect of 1% of the total. Such an evolution of heat could result from induction heating of the film during a slow approach to the equilibrium distribution of trapped flux. In many cases during the experiments, a small effect with a time constant on the order of minutes was noticed in addition to the ordinary thermal relaxation time constant, when the field was increased. This effect did not appear upon decreasing the field or upon increasing H from zero to a value previously exceeded. If indeed this is related to a slow approach to equilibrium in the flux distribution and associated induction heating, the sign of the bridge unbalance produced by the heating would give an apparent depression of $\Delta\kappa$, consistent with the sign of the observed effect.

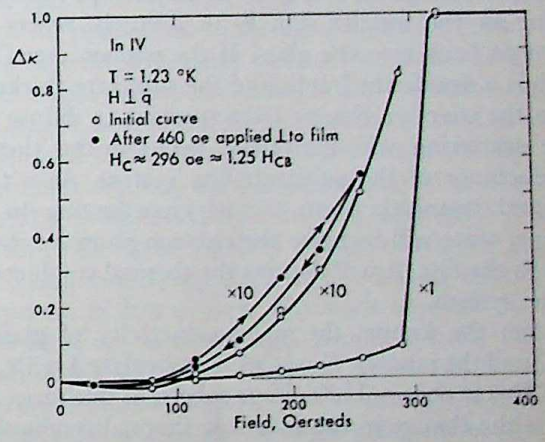


FIG. 10. Hysteresis effects observed in a thick superconducting film.

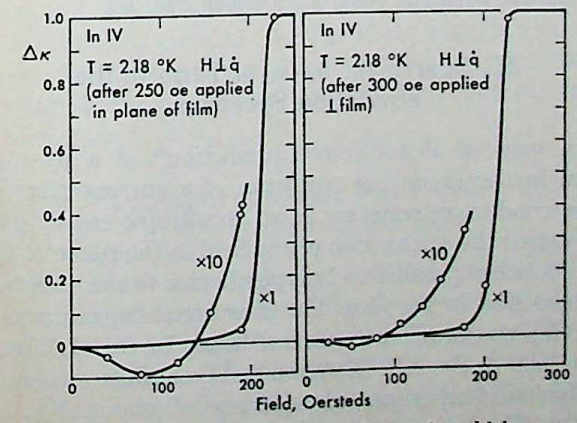


FIG. 9. Change of thermal conductivity of a thick superconducting film with magnetic field applied perpendicular to $\dot{\mathbf{q}}$, the direction of heat flow. At the right, the effect of prior exposure of sample to a large field perpendicular to the film is shown.

Another possible explanation of the long time constant term is heat released during adsorption of helium gas on the thermometer or sample surface when the equilibrium temperature of the heater-thermometer changes. It would be difficult, however, to explain the absence of the effect with decreasing field on this basis.

Yet another explanation might be depression of phonon conductivity by scattering from excitations as the gap decreases. It may be easily shown that the phonon conductivity of the experimental films, if we neglect the substrate, is very small in comparison with the electronic contribution.[27] However, diffuse scattering of the substrate phonons by the electronic excitations in the film when the substrate phonons cross the boundary into the film may lead to an effective "modu-

[27] $K_{\text{electron}} = \frac{1}{3} C_e v_F \Lambda_e$ while $K_{\text{phonon}} = \frac{1}{3} C_p v_p \Lambda_p$. C_e and C_p are comparable in the range 1–4°K in lead, tin, and indium, while $v_F/v_p \approx 10^2$. If $\Lambda_e \approx \Lambda_p \approx d$, then $K_p/K_e \approx 10^{-2}$.

CC-0. In Public Domain. Panini Kanya Maha Vidyalaya Collection.

Digitized by Arya Samaj Foundation Chennai and eGangotri

lation" of the phonon conductivity of the substrate. The resultant effect may be comparable to that of the electronic thermal conductivity of the film. This possibility is examined in the following section.

F. Modulation of Phonon Conductivity of Dielectric Substrate

The wavelength of thermal phonons in glass at 1°K is approximately 2000 Å. Thus we may expect almost perfect specular reflection at the optically polished surfaces of our film substrates. Thermal phonons at 1°K have $\nu \approx 1/400$ of the Debye cutoff frequency, and therefore behave essentially as ordinary ultrasonic phonons. Experiment and theory indicate that the acoustical match between quartz and indium is quite good and leads to less than 1% reflection. Therefore, thermal phonons from the glass substrate pass quite freely into an indium surface film. Neglecting electron-phonon scattering in the film, the phonons will be specularly reflected at the smooth film-vacuum surface, and pass back into the glass. The effect of electron-phonon scattering in the indium film is to diffusely reflect the phonons back into the glass. If the phonon mean free path is a significant fraction of the substrate thickness, then the effective change from specular to diffuse surface scattering will materially decrease the thermal conductance of the substrate-film system. As a field-induced transition from the superconducting to the normal state will increase the electron-phonon interaction in the film, it will depress the thermal conductance of the system.

From the known thermal conductivity of glass at 2°K, and the relation $K=\frac{1}{3}Cv\Lambda$, we calculate $\Lambda \approx 5\times10^{-4}$ cm. This is only $\sim 1/20$ of our substrate thickness, and limits the change in the substrate thermal conductance which may be produced by surface scattering to less than $\sim 5\%$.

The degree of diffuse phonon scattering in the film depends, of course, upon the size of Λ_{pe}, the phonon mean free path as limited by electron scattering, compared with the film thickness. We may estimate Λ_{pe} by using the principle of detailed balance and the measured values of K_{ep}, the phonon-scattering limited electronic thermal conductivity, and its temperature dependence. Kittel[19] has done this in the high-temperature limit. Carrying out a similar analysis in the low-temperature limit, one finds $\Lambda_{pe} \approx 10^{-5}$ cm for tin and 10^{-6} cm for lead at 1.2°K. Since our experimental films range from 5×10^{-6} to 4×10^{-5} cm in thickness, substantial diffuse scattering of phonons may occur within the film in the normal state. When the film becomes superconducting, this diffuse scattering will nearly vanish, since most of the electrons in a superconductor at temperatures far below T_c are in the superconducting ground state and thus unable to scatter phonons. Even though the total modulation of the conductance of the substrate is limited to 5% in our case, since the substrate conductance is typically 2 to 10 times that of the film, the partial modulation obtained at moderate fields might well be enough to account for the small decrease in conductance at low fields found for In II (Fig. 5). By employing a substrate which has a larger mean free path than glass, the effect discussed in this section might be made quite large. Use of a material such as crystalline quartz (with Λ of several cm at 2°K) would appear to permit complete dominance of the thermal conductance by surface effects.

G. Hysteresis Effects in a Thick Film

During the measurements made on In IV ($d=8\lambda$) rather peculiar hysteresis effects were observed (see Figs. 9 and 10). These effects appeared only when $\mathbf{H}$ was perpendicular to $\dot{\mathbf{q}}$, the direction of heat flow. There was no detectable hysteresis when $\mathbf{H}$ was parallel to $\dot{\mathbf{q}}$, nor was any comparable hysteresis observed in any experiments with the other (thinner) experimental films. We consider the most reasonable explanation to be scattering of the heat-carrying electronic excitations at normal-superconducting boundaries.

We may assume that field inhomogeneity and film curvature make necessary small normal-state strips to carry flux through the film. If the super and normal regions have the shape of narrow lamina lying parallel to the direction of $\mathbf{H}$, then when $\mathbf{H}\perp\dot{\mathbf{q}}$ a large number of interfaces would obstruct the heat flow when a moderate field is applied.[16] Hysteresis would be a result of trapped flux affecting the arrangement of domains. Saturation of the film with trapped flux by application of a large field perpendicular to the film surface before a measurement will ensure that the maximum amount of scattering at S-N boundaries occurs before the measurement field is applied. This procedure does appear to suppress the initial decrease of $K(H)$, as we see in Fig. 9. Following saturation of the film with trapped flux, application of a subcritical field parallel to the plane of the film would sweep out some of the previously trapped flux and partially restore the original behavior, consistent with the results shown in Fig. 10.

H. Search for $p_F \cdot v_{drift}$ Term in the Excitation Spectrum

In view of Bogoliubov's prediction[14] of a $\mathbf{p}_F \cdot \mathbf{v}_{drift}$ term in the excitation spectrum of a current-carrying superconductor, resulting in an anisotropic energy gap, we varied the orientation of the field in the plane of the film between parallel and perpendicular to the direction of heat flow in the film. The directional dependence of $\Delta K(H)$ was studied in several films. In tin II, Fig. 6 ($d=2600$ Å, $T=2.12$°K, $\Lambda_e \approx 600$ Å), no effect was found within the experimental accuracy of about 1%. In indium II, as may be seen in Fig. 5, the result is not so clear-cut. Nevertheless, the apparent deviation of the $\mathbf{H}\perp\dot{\mathbf{q}}$ points from those for $\mathbf{H}\|\dot{\mathbf{q}}$ is not of significant

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

magnitude, and the deviations may be explained by poor control of the alignment of $\mathbf{H}$ in the plane of the film during the experiments. An experiment was conducted at 2.1°K on film Pb I, which was similar to Pb II (Fig. 11) except for an unexplained initial negative portion of its $\Delta\kappa(H)$ curve. The curves with $\mathbf{H}\perp\hat{\mathbf{q}}$ and $\mathbf{H}\parallel\hat{\mathbf{q}}$ were identical within the experimental uncertainties, except for a shift in H_c probably arising from a difference in the accuracy of alignment of H in the plane of the film. Since the induced diamagnetic currents are orthogonal to the field, these measurements also show the absence of a dependence on the angle between heat current and diamagnetic current. Thus, if we interpret our results in terms of a decrease of the energy gap with increasing field, the modified gap appears to remain essentially isotropic. Of course, even if there were a $\mathbf{p}\cdot\mathbf{v}$ term, first-order effects would cancel in a transport property such as K, which involves an average of excitations over the entire Fermi surface. However, as the change of ϵ_0 with H approaches 100%, noncancelling second-order effects would be expected to appear, and these would be expected to differ by a factor of several for the parallel and perpendicular cases, because of a different angular average. Thus, another factor must enter to explain the lack of anisotropy. It may well have to do with the fact that in these films, the two equal and opposite surface currents are separated by a distance comparable to ξ_0. Alternately, the short mean free path ($\sim$800 Å) in these samples would be expected to average out anisotropies as in the dirty superconductor theory of Anderson.[28] Both of these characteristics of the experimental situation contrast strongly with the Bogoliubov idealization, which corresponds to a spatially uniform current arising from a displaced sphere of momentum eigenfunctions characteristic of an ideal sample.

The apparent lack of anisotropy is important in evaluating the results of tunnel effect measurements on the magnetic field dependence of the energy gap since, for specular tunneling, this method determines the energy gap normal to the surface of the film, this direction being perpendicular to the directions of both the applied magnetic field and the induced current. This directional aspect of the tunneling experiment would also interfere with any attempt to observe the $\mathbf{v}\cdot\mathbf{p}$ effect in the presence of an externally introduced uniform current flow.

I. Magnetic Fields Not Parallel to the Surface of the Film

We have measured the thermal conductivity of several of our experimental films with field intentionally applied at an angle to the surface. The results were quite striking and unexpected.

Tin II and indium IV are films thick enough to show first-order transitions in fields parallel to the surface; that is, the energy gap is only slightly depressed from its zero field value even just below the transition. This may be seen by inspection of Fig. 4. Application of a field perpendicular to the film might be expected to produce an intermediate state configuration with the fraction of the material in the normal state approximately given by H/H_c.[29] Thus, we might expect $\Delta\kappa(H)$ to increase linearly from zero at $H=0$ to 1.0 at $H=H_c$, although the curve might "sag" somewhat as a result of both additional scattering of the electrons at S-N boundaries and the partial flux penetration of the superconducting regions.

Our measurements contrast strikingly with the above prediction. As we see in the right-hand portions of Figs. 6–8, $\Delta\kappa(H)$ apparently increases exactly as H^2 up to the normal-state value, which is reached in these samples at a field between 42% and 55% of the bulk critical field. Our results suggest that the transition is essentially complete when $K(H)$ first reaches K_n, as there is no change in K with further increase of H (see Fig. 7).

A possible explanation of the above behavior lies in the fact that for these films d is comparable with λ, ξ_0, and Λ_e, in contrast to the usual case of a thick plate in the intermediate state with a transverse field. In the latter case the superconducting and normal domains are arranged as thin lamina parallel to the field in order to minimize the free energy.[16] This is not possible in our thin film samples, since domains shaped as thin lamina parallel to the field would require adjacent S-N boundaries to be closer together than a coherence length ξ_0, and this would contribute an excessively large positive term to the free energy. We might suppose that, when our films are in a moderate magnetic field, the superconducting domains are blunt lamina whose width is determined so as to minimize the sum of the free energy terms arising from the demagnetizing factor of a blunt superconducting domain and from close spacing of the S-N domain boundaries.[30] On general grounds, one might expect an array of vortices[31] about normal lines to be more favorable energetically than lamina involving normal surfaces. Such a structure would have qualitative properties very similar to the laminar structure.

We have measured $\Delta\kappa(H)$ for Pb II with the field applied at angles θ of 0, 1.3 ($\pm$0.4), 5, 10, 20, and 90° with respect to the plane of the film at $T=2.16°$, as

[28] P. W. Anderson, Phys. Chem. Solids 11, 26 (1959).

[29] This is expected because the field at a superconducting-normal interface in a bulk sample must be equal to H_c; then H will be close to H_c throughout the small normal domains, while in the interior of the superconducting domains $H\approx 0$ (if the domains are much larger than the penetration depth λ). Thus, the flux HA which passes through area A of the film penetrates normal regions of total area $A_n=HA/H_c$.

[30] Since our domains will have dimensions of the same order as the coherence length, the energy gap may be nonzero nearly everywhere. In this case we may imagine the energy gap to oscillate as a function of position, with the magnetic flux density being greater in the regions where the gap is small. The remainder of the argument is qualitatively unchanged.

[31] M. Tinkham, Phys. Rev. 129, 2413 (1963).

CC-0. In Public Domain. Panini Kanya Maha Vidyalaya Collection.

Digitized by Arya Samaj Foundation Chennai and eGangotri

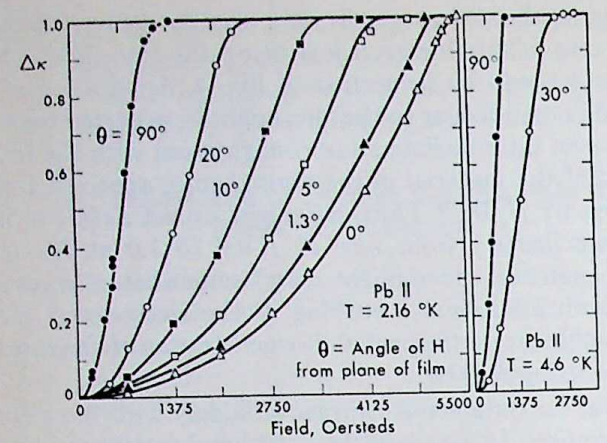


FIG. 11. Change of thermal conductivity of a superconducting lead film with the field applied at various angles with respect to the plane of the film.

well as at 0, 30, and 90° for $T=4.5°$. The results are shown in Fig. 11. We see that as θ is increased from zero the apparent critical field decreases, but the transition appears to remain well defined. The general shape of $\Delta\kappa(H)$ does not change rapidly with θ, but for large θ, $\Delta\kappa(H)$ becomes more nearly linear in H. At $\theta=90°$, the apparent critical field has decreased to a value of the order of H_{cb}.

An attempt has been made to fit the experimental values of $H_c(\theta)$ with various analytical interpolation functions. These are based on the assumption that the parallel and perpendicular components of the magnetic field produce additive contributions to the free energy, and that when the sum of these effects reaches a critical value the film is driven normal. Assuming the contributions to the energy to increase as simple powers of the components of the field, this leads to a form

$$\left(\frac{H_c\cos\theta}{H_{c\parallel}}\right)^m+\left(\frac{H_c\sin\theta}{H_{c\perp}}\right)^n=1.$$

Since magnetic field energies normally are proportional to H^2, one might expect that $m=n=2$. However, this gives a very poor fit at small angles. In fact, a considerably better fit was found to result from the choice $m=n=1$. Subsequently, it was shown[31] by detailed consideration of the vortex model for the perpendicular field case that one would expect $n=1$, but $m=2$. A comparison of the degree of fit of these three interpolation formulas to the critical field data on Pb II has been given previously.[31] A similar comparison with data of White on a thin indium film has also been given.[32] Both of these comparisons favor the theoretically based function, with $m=2$ and $n=1$.

J. Behavior of Thin Lead Films

As we have discussed previously, the predictions of the GLG theory fail at low temperatures. The low-temperature limiting behavior of $\epsilon_0(H)/\epsilon_0(0)$ would be of interest in guiding theoretical efforts. Since facilities for producing stabilized temperatures below 1.2°K were not available to us, experiments were undertaken on lead since 1.2°K equals only 0.18 T_c for this metal. Our detailed results are not easy to interpret and may be a reflection of the peculiarities of superconducting lead, rather than characteristic of superconductors in general. However, we may assert with some confidence that the superconducting transition in a field is of second order for thin films ($d<\lambda$) over the range of temperatures from 0.63 T_c at least down to 0.18 T_c. This conclusion follows from the observed absence of any discontinuity in $\Delta\kappa$ at H_c, as was discussed in Sec. IIIC.

Experimental results for Pb II, a lead film about 500 Å thick, are shown in Fig. 12. Another film,[33] Pb I, gave similar results except for an unexplained initial decrease in $\Delta\kappa$ which appeared at 4.2, 2.1, and 1.2°. The initial decrease of $\Delta\kappa$ for Pb I at 2.1°K was the same for $\mathbf{H}\perp\dot{\mathbf{q}}$ and $\mathbf{H}\|\dot{\mathbf{q}}$. Because of this divergence of results between Pb I and Pb II, we have less confidence in the data in the low-field region than we have near H_c. In Fig. 12, the rounded approach of the 4.5° data to the normal-state conductivity may be accounted for quantitatively by the 0.6° temperature drop along the film, and the consequent 15% critical field spread (Table I). In contrast, at 2.16 and 1.27° the critical field spreads are small (see Table I), and the rounding of the curves at the transition results from the form for small ϵ_0/kT of the G function (Fig. 3) connecting K and ϵ_0/kT, as discussed in the analysis of the In III data.

If we naively invert the data of Fig. 12 point-by-point in the usual manner we find the apparent relations between ϵ_0 and H which are given in Fig. 13 for the various temperatures. At each temperature, the upper curve follows from assuming the zero-field gap

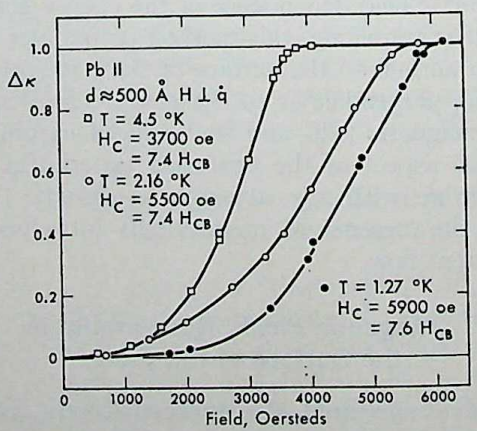


FIG. 12. Change of thermal conductivity of a superconducting film with magnetic field at several temperatures.

[32] M. Tinkham, Rev. Mod. Phys. **36**, 268 (1964).

[33] This film was prepared in a similar manner to Pb II, and has a similar critical field, but it was exposed to the atmosphere for 5 h before mounting, compared to an exposure of ½ h for Pb II. However, we expect that this duration of exposure would have produced negligible oxidation (Sec. IIF).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

value based on the BCS result $\epsilon_0=1.75\ kT_c$ at $T=0°K$. If we take instead the value 2.1 kT_c as suggested by several experiments[34–36] the lower curves result.

If one were to take Fig. 13 at face value, one would be forced to conclude that the variation of ϵ_0 with H is different in lead and indium at the same reduced temperature $t(=T/T_c)$. Although the 4.5° lead curve and 2.1° indium curve (Fig. 4), both with $t\approx\frac{2}{3}$, seem to be consistent with the GLG result $\epsilon_0(H)/\epsilon_0(0)=[1-(H/H_c)^2]^{1/2}$, we see that lead at 2.16° seems to be incompatible with indium at 1.2° even though the reduced temperatures are nearly the same ($t\approx0.35$). It is just possible that the experimental curves for Pb II reflect some spurious effect of the sort that produced the initial negative curve for Pb I, although the deviation would have to be of opposite sign from that observed in Pb I in order to make the Pb II and In III results compatible. Inversion of the Pb I data produces an $\epsilon_0(H)/\epsilon_0(0)$ versus H^2 curve with the same qualitative deviation from the indium curve, once the field is greater than about $(H/H_c)^2=0.3$, [so that we are well above the region where $K(H)$ decreases from the zero-field value].

An alternative, and probably more sound, explanation of the discrepancy is that the relation derived from BRT which we have used to connect the thermal conductivity variations with changes in the energy gap is not completely valid, at least in the case of lead.

It is well known that the thermal and electrodynamic properties of lead and mercury deviate considerably from the law of corresponding states. It has not been possible to accurately test the applicability of Sec. 3 of the BRT theory to these metals, because phonon scattering contributes to the thermal resistance as a result of the low values of θ_{Debye}. But, since $K_e=\frac{1}{3}C_e v_F\Lambda_e$ and the experimentally determined[37] $C_e(T)$ is different from that predicted by BCS, on which the BRT theory is based, we may expect that the BRT calculation will not account satisfactorily for the thermal conductivity of lead and mercury.

Upon comparison of Figs. 4 and 13, we note that if the effective value of the energy gap of lead were to decrease from the usual value at 4.5° to about 1.4 kT_c at 2.16° and about 1.0 kT_c at 1.27°, both of the low curves for lead could be made to coincide with the 1.3° indium III curve. This suggests that we might combine the $\Delta\kappa(H)$ curves of Pb II with the $\epsilon_0(H)/\epsilon_0(0)$ result found for In III at 1.3°, and find the function $G_1(x)$ which would be necessary to connect them. If we now subtract $G(x)$ from $G_1(x)$, we find that the difference, $\Delta G(x)$, is a function which is zero at $x=0$, increases to a maximum of about 0.35 for x between 4 and 6, and then decreases with further increase of x. Let us generalize $G(x)$ to include an "effective density of states"

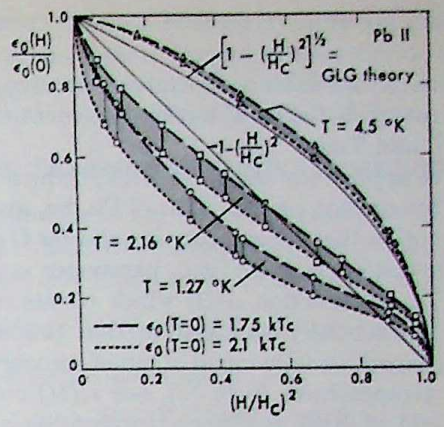


FIG. 13. Dependence of energy gap of a lead film on field as calculated from the data of Fig. 12 using the $G(\epsilon_0/kT)$ function from Fig. 3. For each experimental point, two gaps are computed as shown: one assuming the gap ($2\epsilon_0$) approaches 3.5 kT_c at $T=0$ as found by BCS; and one assuming the gap goes to 4.2 kT_c as found experimentally.

$\rho[E/\epsilon_0(H)]$. Note that this brings us completely outside the BRT theory, and thus we can have no confidence in the result except as a phenomenological description. Let

$$G_1(x)=\int_0^\infty \rho(E/\epsilon_0)E^2(\partial f/\partial E)dE \Big/ \int_0^\infty E^2(\partial f/\partial E)dE$$

and take $\rho(E/\epsilon_0)=1$ for $E\geq\epsilon_0$ (the usual BRT form). Then

$$\Delta G(x)=\int_0^{\epsilon_0} \rho(E/\epsilon_0)E^2(\partial f/\partial E)dE \Big/ \int_0^\infty E^2(\partial f/\partial E)dE.$$

We have attempted to fit $\Delta G(x)$ with several simple forms of $\rho(E/\epsilon_0)$. Within our very limited accuracy of ΔG, we find a satisfactory fit for the delta function $\rho(E/\epsilon_0)=\frac{2}{3}\delta[(E/\epsilon_0)-\frac{1}{2}]$, or for the parabolic function $\rho(E/\epsilon_0)=\frac{3}{2}(E/\epsilon_0)^2$. We emphasize that we do not take the above to be conclusive evidence for the existence of "states in the gap" of lead, or for the location of any such hypothetical states. Rather, these considerations are intended to indicate the major scale of the departure of the data on lead from that on indium at low reduced temperatures. We conclude that lead behaves in some respects differently from other superconductors even at the same reduced temperatures, but the cause is not at present fully understood.

IV. CONCLUSIONS

From the results of the preceeding sections we may draw several conclusions: First, the thermodynamic transition in a field may be quite sharp and well-defined even for films sufficiently thin (700 Å) that $H_c>7\,H_{cb}$. The thermodynamic transition appears to be reversible and without hysteresis for typical films up to 2500 Å thick. Our experiments give direct evidence that the superconducting transition is of second order in

[34] P. L. Richards and M. Tinkham, Phys. Rev. 119, 575 (1960).
[35] D. M. Ginsberg and M. Tinkham, Phys. Rev. 118, 990 (1960).
[36] I. Giaever, Phys. Rev. Letters 5, 147 (1960).
[37] D. L. Decker, D. E. Mapother, and R. W. Shaw, Phys. Rev. 112, 1888 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection

Digitized by Arya Samaj Foundation Chennai and eGangotri

films which are thin (compared with the penetration depth) for temperatures ranging from T_c down to 0.18 T_c, the lowest available temperature, in disagreement with the theory of Bardeen[8] but in agreement with that of Nambu and Tuan.[9]

We have applied the theory of BRT,[13] which is based on BCS, to our data on thin films of Pb, Sn, and In, on the assumption that their theory is valid for $H \neq 0$ with a suitable field-dependent gap parameter $\epsilon_0(H)$. At $T=0.65\ T_c$, the function $\epsilon_0(H)$ which results from our analysis is essentially identical to that predicted by GLG and experimentally found in tunnel measurements. At lower temperatures ($0.36\ T_c$), our $\epsilon_0(H)$ computed with the aid of BRT is different, indicating a breakdown of the GLG theory for $T \ll T_c$, or possibly a failure of the method of data analysis via BRT.

By varying the relative orientation of the direction of heat flow and the applied magnetic field in the plane of the experimental films, we unsuccessfully sought evidence for vestiges of the $\mathbf{p}_F \cdot \mathbf{v}$ term in the excitation spectrum of a current-carrying superconductor, predicted by Bogoliubov.[14] Since the experimental conditions differed radically from those envisioned by Bogoliubov, this negative result is not surprising.

Application of magnetic fields at various angles up to 90° with respect to the film plane still produced surprisingly well defined thermodynamic transitions, although H_c was reduced considerably below the parallel field value. The transition in a perpendicular field appeared to be second-order even for films of thickness up to 4000 Å, whose parallel field transition was definitely first order. The perpendicular critical field values and the angular dependence of H_c seem to be explained satisfactorily by the theory of Tinkham.[31]

Lead films gave results qualitatively similar to indium and tin films, but quantitatively the form of the derived $\epsilon_0(H)$ at low temperatures was different from that of tin and indium. We suspect that this is connected with the well-known peculiarities in the thermal properties of superconducting lead.

ACKNOWLEDGMENTS

Thanks are due to I. F. Silvera for measuring the effect of atmospheric oxidation on the optical transmissivity of our Pb films. One of us (D.M.) would like to thank Dr. K. G. Ramanathan for the hospitality of the National Physical Laboratory of India during the preparation of the manuscript.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Higher Order Coherent Raman Effects*

C. L. Tang
Raytheon Research Division, Waltham, Massachusetts
(Received 8 January 1964)

With the recently available very intense coherent maser light, interesting higher order coherent Raman effects have been observed in solids, liquids, and gases. Garmire, Pandarese, and Townes have given a phenomenological theory of these higher order coherent Raman effects. In this paper a somewhat more detailed theory of these effects is given. The analysis is based on a perturbation solution of the Boltzmann equation for the density matrix of the Raman active medium. General expressions for all the relevant induced dipole moments and absorbed or emitted Stokes and anti-Stokes radiations in the ordinary or higher order coherent Raman processes are given. These cover a great variety of possible experimental situations. Several specific simple examples are studied in some detail and the results of the phenomenological theory are reproduced from the general results given here.

I. INTRODUCTION

THE first-order Raman effects have been investigated extensively in the past both experimentally and theoretically. With the recently available very intense maser light, interesting higher order coherent Raman effects in solids,[1] liquids,[2] and gases[3] have also been observed. Garmire, Pandarese, and Townes[4] give a detailed macroscopic theory of these higher order coherent Raman effects, particularly those through excitation of intense coherent molecular oscillations at infrared frequencies. Many interesting conclusions were drawn from the theory and some of these have already been observed experimentally. In this paper an attempt is made to give a somewhat more detailed analysis of the same effects on the basis of the quantum-mechanical Boltzmann equation for the density matrix characterizing the Raman-active medium and to include in the theory the effects of various relaxation processes.

* Work supported in part by Contract No. AF 19(628)-3862, AFCRL, Hanscom Field, Bedford, Massachusetts.

[1] G. Eckardt, D. P. Bortfeld, and M. Geller, Appl. Phys. Letters 3, 137 (1963).

[2] B. Stoicheff, Symposium on Molecular Structure and Spectroscopy, Ohio State University, June 1963 (unpublished); Phys. Letters 7, 186 (1963); H. J. Zeiger, P. E. Tannenwald, S. Kern, and R. Herendeen, Phys. Rev. Letters 11, 419 (1963).

[3] R. W. Mink, R. W. Terhune, and W. G. Rado, Appl. Phys. Letters 3, 181 (1963).

[4] E. Garmire, F. Pandarese, and C. H. Townes, Phys. Rev. Letters 11, 160 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The effects of spatial dispersion[5] are not, however, explicitly taken into account in the calculations here. This may be a serious omission in some cases involving liquids and solids, but in the cases of gases it is most likely not. When spatial dispersion effects are neglected, a local description of the Raman-active dielectric medium is possible. The study of the various higher order coherent Raman effects can then be carried out in several steps.

In the first step, we consider a volume of the dielectric with linear dimensions large compared with the intermolecular spacings but small compared with the wavelengths of the interacting photons and the coherent molecular vibrational waves (or other polarization waves). Within a volume of this size, the radiation fields can be assumed coherent and uniform, and the medium can be characterized by a spatially independent density matrix. The quantum-mechanical Boltzmann equation governs the time response of the density matrix due to the applied fields. The average induced dipole moment per molecule or molecular oscillations in the medium produced in various Raman processes can be obtained from the corresponding higher order terms in the perturbation solution of the Boltzmann equation. From these one can determine the frequencies of the Raman-scattered Stokes and anti-Stokes radiations and the rate at which they are absorbed or emitted locally. In Sec. II an explicit formulation of the problem in terms of the density matrix is given along with a formal procedure for obtaining the appropriate perturbation solution.

In the second step, we consider a volume of the dielectric medium with linear dimensions large compared with the wavelengths of all the waves of interest but still small compared with the distances over which the amplitudes and phases of the waves change appreciably. The harmonic spatial dependences of all the waves can therefore be directly introduced into the various expressions obtained in the previous step. When macroscopic spatially coherent effects are considered, one assumes a fixed phase for each of the electromagnetic waves. Upon averaging the resulting expressions over such a volume, one obtains the conditions under which each Stokes or anti-Stokes radiation will be absorbed or amplified on the average. These include the so-called phase-matching conditions which give the propagation vectors of the Stokes and anti-Stokes radiations.

In the final step, one may consider questions which are only significant on dimensional scales over which the amplitudes and phases of the various waves vary appreciably. These involve questions such as energy transport, boundary effects, geometrical effects, etc.

In the present study we do not, however, go beyond the second step in any detail. The primary concern here is in the intrinsic properties of the higher order coherent Raman processes as discussed by Garmire, Pandarese, and Townes.

The above assumptions and basic steps are implicit in the calculations and discussions given below.

[5] See, for example, J. J. Hopfield and D. G. Thomas, Phys. Rev. 132, 563 (1963).

II. THE DENSITY-MATRIX EQUATION

We ignore first the spatial dependence in all the variables. Under the usual assumptions,[6] the quantum-mechanical Boltzmann equation for the density matrix is

$$\partial\rho/\partial t=-(i/\hbar)[H,\rho]-\tfrac{1}{2}[\Gamma(\rho-\bar{\rho})+(\rho-\bar{\rho})\Gamma], \quad (2.1)$$

where the total Hamiltonian

$$H=H_0+V=H_0-\mathbf{p}\cdot\mathbf{E} \quad (2.1a)$$

is the sum of the unperturbed Hamiltonian of the molecule, H_0, and the interaction energy of the molecules with the total electric field, $\mathbf{E}$. In a representation in which H_0 is diagonal, Γ, which characterizes phenomenologically the effects of various relaxation processes, is a diagonal matrix $\Gamma_{mn}=\gamma_m\delta_{mn}\cdot 1$, where 1 is a unit matrix with the dimensionality of the corresponding eigenvalue of H_0. The matrix $\bar{\rho}$ is diagonal, and the diagonal elements are proportional to the thermal equilibrium populations of the corresponding energy levels at a given temperature T in the absence of V, or

$$\bar{\rho}=\exp(-H_0/kT)/\mathrm{Trace}[\exp(-H_0/kT)]. \quad (2.1b)$$

$\rho(t)$ determines the average induced electric dipole moment, $\mathbf{p}(t)$, per molecule:

$$\mathbf{p}(t)=\mathrm{Tr}[\mathbf{p}\rho(t)], \quad (2.2)$$

where $\mathbf{p}$ is the electric dipole operator.

In studying Raman effects we consider three energy levels, ϵ_1, ϵ_2, and ϵ_3, of the molecules (Fig. 1) where level 1 is the ground state and level 2 could be a vibrational level, a rotational level, or an electronic level if electronic Raman effects are considered. Level 3 is the intermediate state. All these levels will in general be degenerate. For the basic first-order Raman process, it is sufficient to assume that the total electric field consists of two components:

$$\mathbf{E}(t)=[\mathbf{E}_a e^{i\omega_a t}+\mathbf{E}_a^* e^{-i\omega_a t}]+[\mathbf{E}_b e^{i\omega_b t}+\mathbf{E}_b^* e^{-i\omega_b t}]. \quad (2.3a)$$

The elements of interest in the perturbation term V of the total Hamiltonian are assumed to be:

$$\begin{aligned} &V_{1\alpha,1\beta}=V_{2\alpha,2\beta}=V_{3\alpha,3\beta}=V_{1\alpha,2\beta}=0,\\ &V_{1\alpha,3\beta}=V_{1\alpha,3\beta}^{(a)}e^{i\omega_a t}+V_{1\alpha,3\beta}^{(a)*}e^{-i\omega_a t}, \quad (2.3b)\\ &V_{2\alpha,3\beta}=V_{2\alpha,3\beta}^{(b)}e^{i\omega_b t}+V_{2\alpha,3\beta}^{(b)*}e^{-i\omega_b t}, \end{aligned}$$

for ω_a and ω_b sufficiently close to the correspond-

[6] R. Karplus and J. Schwinger, Phys. Rev. 73, 1020 (1948); W. E. Lamb, Jr., *Lectures in Theoretical Physics* (Interscience Publishers, Inc., New York, 1960), Vol. II, p. 435; Third International Symposium on Quantum Electronics, Paris, France, 1963 (unpublished).

CC-0. In Public Domain. Gurukul K... Collection.

Digitized by Arya Samaj Foundation Chennai and eGangotri

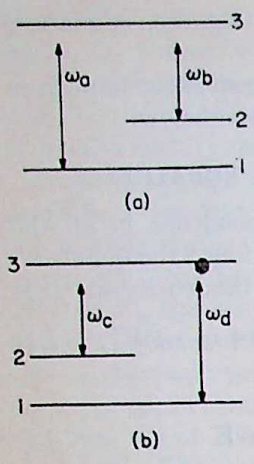


FIG. 1. Schematic diagram of energy levels and fields involved in (a) ordinary and (a), (b) higher order Raman processes.

ing transition frequencies $\omega_{31}=\hbar^{-1}(\epsilon_3-\epsilon_1)$ and $\omega_{32}=\hbar^{-1}(\epsilon_3-\epsilon_2)$, and

$$V_{m\alpha,n\beta}^{(a,b)}=-\langle m,\alpha|\mathbf{p}\cdot\mathbf{E}_{a,b}|n,\beta\rangle. \quad (2.3c)$$

In Eqs. (2.3b) and (2.3c), α and β designate the degenerate states of the energy levels ϵ_m and ϵ_n, respectively. We also assume that

$$\omega_a-\omega_b\approx\hbar^{-1}(\epsilon_2-\epsilon_1)=\omega_{21}, \quad (2.3d)$$

but ω_a and ω_b are not so near the corresponding transition frequencies that the direct processes become important.

As in the usual theory[7] of ordinary Raman effects for a single molecule, which is based on the perturbation solution of Schrödinger's equation, we obtain the steady-state solution of the Boltzmann equation (2.1) by perturbation techniques.[8] Specifically, we formally expand the elements of the density matrix:

$$\rho_{m\alpha,n\beta}=\rho_{m\alpha,n\beta}^{(0)}+\rho_{m\alpha,n\beta}^{(1)}+\cdots+\rho_{m\alpha,n\beta}^{(j)}+\cdots \quad (2.4)$$

in successive orders of approximation corresponding to ascending powers of the amplitude of the electric field. The lowest order term

$$\rho_{m\alpha,n\beta}^{(0)}=\frac{1}{T_{mn}}\int^t\exp\left(i\omega_{mn}+\frac{1}{T_{mn}}\right)\times(t'-t)\bar{\rho}_{m\alpha,n\beta}dt', \quad (2.5)$$

which would simply be equal to $\bar{\rho}_{m\alpha,n\beta}$ if the latter is purely diagonal and time-independent. In Eq. (2.5)

$$T_{mn}^{-1}=\tfrac{1}{2}(\gamma_m+\gamma_n), \quad (2.6)$$

which is independent of the degeneracy indices α and β.

The successive terms, i.e., $j\geq 1$:

$$\rho_{m\alpha,n\beta}^{(j)}=-\frac{i}{\hbar}\sum_{l,\gamma}\int^t\exp\left(i\omega_{mn}+\frac{1}{T_{mn}}\right)(t'-t)\times[V_{m\alpha,l\gamma}\rho_{l\gamma,n\beta}^{(j-1)}-\rho_{m\alpha,l\gamma}^{(j-1)}V_{l\gamma,n\beta}]dt'. \quad (2.7)$$

This series solution, (2.4)–(2.7), is formally identical to that used in the study of nonlinear interaction of oscillating modes in a gas maser, or in similar problems, where the nonlinearity is due to multiple successive single-photon processes.[6,8,9] Raman processes are true multiple-photon processes.

Mathematically, the distinction is quite simple, and it lies in how the intermediate steps in the iterations implied in (2.7) are carried out, apart from the resonance conditions. Very briefly, in the successive single-photon process, one calculates alternately the off-diagonal and the diagonal elements in the successive steps of (2.7) (see Ref. 8). In the ordinary Raman process, as in the situation depicted in Fig. 1(a), in which simultaneous absorption of a photon at ω_a and emission of a photon at ω_b occur, one calculates, for example, the second-order term $\rho_{1\alpha,2\beta}^{(2)}$ directly from $\rho_{1\alpha,3\gamma}^{(1)}$ and $\rho_{3\gamma,2\beta}^{(1)}$. This will become clearer in the following sections.

III. ORDINARY COHERENT RAMAN EFFECTS

In the ordinary Raman process, only two photons are involved as shown in Fig. 1(a). Following the usual procedure,[8] we obtain from Eqs. (2.7), (2.5), (2.3b), and (2.1b) the lowest order terms in the perturbation expansions of the off-diagonal elements of the density matrix

$$\rho_{1\alpha,3\gamma}^{(1)}=-\frac{i\Delta\bar{\rho}_{31}T_{13}V_{1\alpha,3\gamma}^{(a)}e^{i\omega_a t}}{\hbar[1+i(\omega_a-\omega_{31})T_{13}]}, \quad (3.1a)$$

$$\rho_{2\beta,3\gamma}^{(1)}=-\frac{i\Delta\bar{\rho}_{32}T_{23}V_{2\beta,3\gamma}^{(b)}e^{i\omega_b t}}{\hbar[1+i(\omega_b-\omega_{32})T_{23}]}, \quad (3.1b)$$

where

$$\Delta\bar{\rho}_{mn}\equiv\bar{\rho}_{m\alpha,m\alpha}-\bar{\rho}_{n\beta,n\beta}\equiv\bar{\rho}_{mm}-\bar{\rho}_{nn} \quad (3.2)$$

for any α and β. $\Delta\bar{\rho}_{mn}$ is diagonal and is independent of the degeneracy indices by virtue of (2.1b). Substituting (3.1a) and (3.1b) into (2.7), we obtain for the second-order term in the off-diagonal matrix element

$$\rho_{1\alpha,2\beta}^{(2)}=\sum_{\gamma}\frac{V_{1\alpha,3\gamma}^{(a)}V_{3\gamma,2\beta}^{(b)*}T_{12}e^{i(\omega_a-\omega_b)t}}{\hbar^2[1+i(\omega_a-\omega_b-\omega_{21})T_{12}]}\times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1-i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1+i(\omega_a-\omega_{31})T_{13}}\right]. \quad (3.3)$$

In the proper limits this term corresponds to the second-

[7] See, for example, M. Göppert-Mayer, Ann. Phys. 9, 273 (1931); G. Placzek, *Handbuch der Radiologie* (Akademische Verlagsgesellschaft Leipzig, 1934), Vol. VI, Part II, pp. 209–347. For more recent studies in connection with maser work see, for example, A. Javan, J. Phys. Radium 19, 806 (1958).

[8] C. L. Tang and H. Statz, Phys. Rev. 128, 1013 (1962); H. Statz and C. L. Tang, Third International Symposium on Quantum Electronics, Paris, France, 1963 (unpublished).

[9] B. Senitzky, G. Gould, and S. Cutler, Phys. Rev. 130, 1460 (1963); H. Haken and H. Sauerman, Z. Physik 173, 261 (1963); N. Bloembergen and Y. R. Shen, Phys. Rev. 133, A37 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

order term in the perturbed wave function of the Schrödinger equation in the theory of ordinary Raman effects for a single molecule.[7] It contains all the necessary information for calculating the coherent molecular oscillations excited by the light waves. If electronic Raman processes are being considered and dipole transitions between levels 1 and 2 are allowed, it describes the induced dipole moment oscillating at $\omega_{21}=\omega_a-\omega_b$ produced by the light waves at ω_a and ω_b.

This coherently driven molecular oscillation described by $\rho_{1\alpha,2\beta}^{(2)}$ gives rise to a third-order term $\rho_{1\alpha,3\gamma}^{(3)}$, which results in an oscillating dipole moment per molecule at ω_a:

$$\mathbf{p}_a^{(3)}=\sum_{\alpha,\gamma}[\mathbf{p}_{3\gamma,1\alpha}\rho_{1\alpha,3\gamma}^{(3)}+\rho_{3\gamma,1\alpha}^{(3)}\mathbf{p}_{1\alpha,3\gamma}]$$
$$=\sum_{\gamma,\alpha,\gamma',\beta}i\chi_a\mathbf{p}_{3\gamma,1\alpha}V_{1\alpha,3\gamma'}^{(a)}V_{3\gamma',2\beta}^{(b)*}V_{2\beta,3\gamma}^{(b)}\times e^{i\omega_a t}+\text{c.c.}, \quad (3.4a)$$

where

$$\chi_a=\frac{T_{12}T_{13}}{\hbar^3[1+i(\omega_a-\omega_b-\omega_{21})T_{12}][1+i(\omega_a-\omega_{31})T_{13}]}\times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1-i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1+i(\omega_a-\omega_{31})T_{13}}\right], \quad (3.4b)$$

and an oscillating dipole moment at ω_b:

$$\mathbf{p}_b^{(3)}=\sum_{\beta,\gamma}[\mathbf{p}_{3\gamma,2\beta}\rho_{2\beta,3\gamma}^{(3)}+\rho_{3\gamma,2\beta}^{(3)}\mathbf{p}_{2\beta,3\gamma}]$$
$$=\sum_{\gamma,\beta,\gamma',\alpha}i\chi_b\mathbf{p}_{3\gamma,2\beta}V_{2\beta,3\gamma'}^{(b)}V_{3\gamma',1\alpha}^{(a)*}V_{1\alpha,3\gamma}^{(a)}\times e^{i\omega_b t}+\text{c.c.}, \quad (3.4c)$$

where

$$\chi_b=\frac{T_{12}T_{23}}{\hbar^3[1-i(\omega_a-\omega_b-\omega_{21})T_{12}][1+i(\omega_b-\omega_{32})T_{23}]}\times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1+i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1-i(\omega_a-\omega_{31})T_{13}}\right], \quad (3.4d)$$

with the help of (3.3), (2.7), (2.3b), and (2.2). Not all the terms in the sums over the degeneracy indices in (3.4a) and (3.4c), as well as (3.3), are nonzero. Many will vanish of course due to selection rules or other special circumstances.

These dipole moments at frequencies ω_a and ω_b, $\mathbf{p}_a^{(3)}$ and $\mathbf{p}_b^{(3)}$, induced in the Raman process will lead to energy exchange between the molecules and the oscillating electric fields at the corresponding frequencies. The average absorbed power is equal to the time average of the scalar product of the rate of decrease of the dipole moment and the corresponding oscillating component of the electric field. We obtain from (3.4a) for the average power absorbed per molecule at the frequency ω_a

$$P_a=\text{Re}\{\sum_{\gamma,\alpha,\gamma',\beta}2\omega_a\chi_a V_{3\gamma,1\alpha}^{(a)*}V_{1\alpha,3\gamma'}^{(a)}\times V_{3\gamma',2\beta}^{(b)*}V_{2\beta,3\gamma}^{(b)}\}. \quad (3.5)$$

Similarly, from (3.4c) the average absorbed power per molecule at ω_b is

$$P_b=\text{Re}\{\sum_{\gamma,\beta,\gamma',\alpha}2\omega_b\chi_b V_{3\gamma,2\beta}^{(b)*}V_{2\beta,3\gamma'}^{(b)}\times V_{3\gamma',1\alpha}^{(a)*}V_{1\alpha,3\gamma}^{(a)}\}. \quad (3.6)$$

We next introduce the spatial dependence of the perturbation terms $V^{(a)}$ and $V^{(b)}$ into (3.5) and (3.6) and obtain the absorbed power averaged over a volume large compared with all the wavelengths of interest but small compared with the distances over which the amplitudes and phases of the waves vary appreciably. We consider homogeneous plane waves:

$$V_{1\alpha,3\beta}^{(a)}=\bar{V}_{1\alpha,3\beta}^{(a)}e^{-i\mathbf{k}_a\cdot\mathbf{r}},$$
$$V_{2\alpha,3\beta}^{(b)}=\bar{V}_{2\alpha,3\beta}^{(b)}e^{-i\mathbf{k}_b\cdot\mathbf{r}}, \quad (3.7)$$

where $\mathbf{r}$ is the position vector of the average molecules characterized by the density matrix and $\bar{V}_{m\alpha,n\beta}^{(a,b)}$ is spatially independent but complex.

Substituting (3.7) into (3.5) and (3.6) and averaging over $\mathbf{r}$, we obtain the average power per unit volume absorbed at the frequency ω_a:

$$\langle P_a\rangle_{\mathbf{r}}=\text{Re}\{\sum_{\gamma,\alpha,\gamma',\beta}2N\omega_a\chi_a\bar{V}_{3\gamma,1\alpha}^{(a)*}\bar{V}_{1\alpha,3\gamma'}^{(a)}\times\bar{V}_{3\gamma',2\beta}^{(b)*}\bar{V}_{2\beta,3\gamma}^{(b)}\}, \quad (3.8a)$$

and that at ω_b:

$$\langle P_b\rangle_{\mathbf{r}}=\text{Re}\{\sum_{\gamma,\beta,\gamma',\alpha}2N\omega_b\chi_b\bar{V}_{3\gamma,2\beta}^{(b)*}\bar{V}_{2\beta,3\gamma'}^{(b)}\times\bar{V}_{3\gamma',1\alpha}^{(a)*}\bar{V}_{1\alpha,3\gamma}^{(a)}\}, \quad (3.8b)$$

where N is the density of the active molecules coupled to the fields. In both cases the final results are independent of the propagation vectors. Unlike the higher order Raman effects which will be considered in the following section, there are no phase matching conditions here. The average absorptions per unit volume, (3.8a) and (3.8b), are always directly proportional to the local absorptions (3.6a) and (3.6b).

On the other hand, substituting (3.7) into (3.3), we have

$$\rho_{1\alpha,2\beta}^{(2)}=\sum_{\gamma}\frac{\bar{V}_{1\alpha,3\gamma}^{(a)}\bar{V}_{3\gamma,2\beta}^{(b)*}T_{12}}{\hbar^2[1+i(\omega_a-\omega_b-\omega_{21})T_{12}]}\times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1-i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1+i(\omega_a-\omega_{31})T_{13}}\right]\times\exp i[(\omega_a-\omega_b)t-(\mathbf{k}_a-\mathbf{k}_b)\cdot\mathbf{r}], \quad (3.9)$$

which means the molecular-oscillation wave (or other polarization waves) produced in the coherent Raman process has a propagation vector equal to $\mathbf{k}_a-\mathbf{k}_b$.

If ω_a designates the primary light, say the maser light at ω_l, then (3.8b) describes the first Stokes radiation at

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$\omega_{S_1}=\omega_b$:

$$\langle P_{S_1}\rangle_r=\mathrm{Re}\{\sum_{\gamma,\beta,\gamma',\alpha} 2N\omega_{S_1}\chi_{S_1}\bar{V}_{3\gamma,2\beta}{}^{(S_1)*}\bar{V}_{2\beta,3\gamma'}{}^{(S_1)} \times\bar{V}_{3\gamma',1\alpha}{}^{(l)*}\bar{V}_{1\alpha,3\gamma}{}^{(l)}\}, \quad (3.10a)$$

where

$$\chi_{S_1}=\frac{T_{12}T_{23}}{\hbar^3[1-i(\omega_l-\omega_{S_1}-\omega_{21})T_{12}][1+i(\omega_{S_1}-\omega_{32})T_{23}]} \times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1+i(\omega_{S_1}-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1-i(\omega_l-\omega_{31})T_{13}}\right]. \quad (3.10b)$$

If ω_b designates the primary light, say the maser light at ω_l, then (3.8a) describes the anti-Stokes radiation at $\omega_{AS}=\omega_a$:

$$\langle P_{AS}\rangle_r=\mathrm{Re}\{\sum_{\gamma,\alpha,\gamma',\beta} 2N\omega_{AS}\chi_{AS}\bar{V}_{3\gamma,1\alpha}{}^{(AS)*}\bar{V}_{1\alpha,3\gamma'}{}^{(AS)} \times\bar{V}_{3\gamma',2\beta}{}^{(l)*}\bar{V}_{2\beta,3\gamma}{}^{(l)}\}, \quad (3.10c)$$

where

$$\chi_{AS}=\frac{T_{12}T_{13}}{\hbar^3[1+i(\omega_{AS}-\omega_l-\omega_{21})T_{12}][1+i(\omega_{AS}-\omega_{31})T_{13}] \times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1-i(\omega_l-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1+i(\omega_{AS}-\omega_{31})T_{13}}\right]. \quad (3.10d)$$

Equations (3.10a)–(3.10d) contain all the information about the Stokes and anti-Stokes radiations in the ordinary Raman process. These expressions are very general though somewhat complicated. In order to see the physical interpretations of these results better, we consider some simple limiting cases.

We first simplify the χ's. Let us assume that

$$T_{13},T_{23}\to\infty \quad (3.11)$$

and

$$\Delta\bar{\rho}_{32}\approx-\bar{\rho}_{22}, \quad \Delta\bar{\rho}_{31}\approx-\bar{\rho}_{11}. \quad (3.12)$$

Therefore, we have from (3.10b):

$$\lim_{T_{13},T_{23}\to\infty}\chi_{S_1}=\frac{1}{\hbar^3}\left[\frac{T_{12}+iT_{12}{}^2(\omega_l-\omega_{S_1}-\omega_{21})}{1+(\omega_l-\omega_{S_1}-\omega_{21})^2T_{12}{}^2}\right] \times\left[\pi\delta(\omega_{S_1}-\omega_{32})-iP\frac{1}{\omega_{S_1}-\omega_{32}}\right] \times\left[-\bar{\rho}_{22}\pi\delta(\omega_{S_1}-\omega_{32})+i\bar{\rho}_{22}P\frac{1}{\omega_{S_1}-\omega_{32}} -\bar{\rho}_{11}\pi\delta(\omega_l-\omega_{31})-i\bar{\rho}_{11}P\frac{1}{\omega_l-\omega_{31}}\right], \quad (3.13a)$$

where P in (3.13a) designates principal values. The δ-functions' contributions correspond to the effects of the single-photon processes between levels 1 or 2 and 3. These effects are usually not considered as a part of the Raman process. One should note, however, that a clear-cut distinction can easily be made only in the case T_{13}, $T_{23}\to\infty$. Experimentally, in the cases where the Raman processes are important, the direct single-photon processes usually are not, so that one does not have to make a distinction and both the δ-function contributions and the principal value signs in (3.13a) can be omitted:

$$\chi_{S_1}=-\frac{1}{\hbar^3(\omega_{S_1}-\omega_{32})}\left[\frac{T_{12}}{1+(\omega_l-\omega_{S_1}-\omega_{21})^2T_{12}{}^2} +i\frac{(\omega_l-\omega_{S_1}-\omega_{21})T_{12}{}^2}{1+(\omega_l-\omega_{S_1}-\omega_{21})^2T_{12}{}^2}\right] \times\left[\frac{\bar{\rho}_{11}}{\omega_l-\omega_{31}}-\frac{\bar{\rho}_{22}}{\omega_{S_1}-\omega_{32}}\right]. \quad (3.13b)$$

We also assume that the primary light is linearly polarized in, say the $\hat{z}$ direction

$$\mathbf{E}_l=\hat{z}E_le^{-i\mathbf{k}_l\cdot\mathbf{r}}, \quad (3.14a)$$

and that the Stokes radiation is

$$\mathbf{E}_{S_1}=[\sin\theta_{S_1}\hat{x}+\cos\theta_{S_1}\hat{z}]\bar{E}_{S_1} \times\exp(-i\mathbf{k}_{S_1}\cdot\mathbf{r}+i\psi_{S_1}). \quad (3.14b)$$

With (3.13b), (3.14a) and (3.14b), (3.10a) and (3.10b) can be readily evaluated if, for example, levels 1 and 2 are nondegenerate; one obtains for the Stokes radiation the familiar result[4]:

$$\langle P_{S_1}\rangle_r=-\frac{2N(\bar{\rho}_{11}-\bar{\rho}_{22})\omega_{S_1}T_{12}}{9\hbar^3\Delta_1{}^2[1+(\omega_l-\omega_{S_1}-\omega_{21})^2T_{12}{}^2]} \times|p_{13}|^2|p_{23}|^2\bar{E}_l{}^2\bar{E}_{S_1}{}^2\cos^2\theta_{S_1}, \quad (3.15a)$$

where $\Delta_1=\omega_{31}-\omega_l\approx\omega_{32}-\omega_{S_1}$ near the resonance peak at $\omega_l-\omega_{S_1}\approx\omega_{21}$ and $|p_{mn}|$ is the magnitude of the reduced matrix element of the operator $\mathbf{p}$. Since $\bar{\rho}_{11}>\bar{\rho}_{22}$, $\langle P_{S_1}\rangle_r\leq0$, and Stokes radiation is emitted. It is emitted in all directions with the radiation pattern determined by the angle, θ_{S_1}, between the polarizations of the maser light and the Stokes light according to $\cos^2\theta_{S_1}$.

Similarly, one obtains for the absorbed anti-Stokes radiation per unit volume[4]:

$$\langle P_{AS}\rangle_r=\frac{2N(\bar{\rho}_{11}-\bar{\rho}_{22})\omega_{AS}T_{12}}{9\hbar^3\Delta_2{}^2[1+(\omega_l-\omega_{S_1}-\omega_{21})^2T_{12}{}^2]} \times|p_{13}|^2|p_{23}|^2\bar{E}_l{}^2\bar{E}_{AS}{}^2\cos^2\theta_{AS}, \quad (3.15b)$$

where $\Delta_2=\omega_{31}-\omega_{AS}\approx\omega_{32}-\omega_l$. It is clear from (3.15b) that anti-Stokes radiation is always absorbed.

These are only some special cases. A great variety of more complicated situations may arise in experiments and the general expressions (3.10a)–(3.10d) must then be used.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

IV. HIGHER ORDER COHERENT RAMAN PROCESSES

A. General Theory

In the higher order Raman processes more than two photons are involved simultaneously. Let there be photons at angular frequencies ω_c and ω_d as shown in Fig. 1(b) in addition to the photons at ω_a and ω_b as shown in Fig. 1(a).

Physically, a light wave at ω_c, propagating through a medium which is being modulated by coherent molecular oscillations, excited by two other light waves at ω_a and ω_b as described by the second order off-diagonal density matrix element $\rho_{1\alpha,2\beta}^{(2)}$, (3.3), will induce an oscillating dipole moment at $\omega_d=\omega_c+(\omega_a-\omega_b)$. Similarly, the light wave at ω_d can induce an oscillating dipole moment at $\omega_c=\omega_d-(\omega_a-\omega_b)$. Both are higher order Raman effects. This is not unlike the double-resonance effects[10] where the coherence between levels 1 and 2 is induced directly by an applied field of angular frequency ω_{21}. In the present case, this coherence is induced by two applied fields at ω_a and ω_b, where $\omega_a-\omega_b\approx\omega_{31}-\omega_{32}=\omega_{21}$.

Mathematically, the presence of two additional oscillating fields adds new elements in the perturbation term, V, in the total Hamiltonian for the density-matrix equation:

$$V_{1\alpha,3\beta}^{(d)}e^{i\omega_d t}+V_{1\alpha,3\beta}^{(d)*}e^{-i\omega_d t}, \tag{4.1a}$$

and

$$V_{2\alpha,3\beta}^{(c)}e^{i\omega_c t}+V_{2\alpha,3\beta}^{(c)*}e^{-i\omega_c t}. \tag{4.1b}$$

In the perturbation solution of the density-matrix equation, a third-order term in the off-diagonal density-matrix element $\rho_{1\alpha,3\gamma}^{(3)}$ appears which comes from substituting into (2.7) the second-order off-diagonal elements $\rho_{1\alpha,2\beta}^{(2)}$ and $\rho_{2\beta,1\alpha}^{(2)}$ and the new perturbation terms given in (4.1b). Similarly, a third-order term $\rho_{2\beta,3\gamma}^{(3)}$ appears. These give rise to new induced dipole terms.

From (2.2), (2.7), (3.3), and (4.1a), we then obtain the average induced dipole moment per molecule at $\omega_d-(\omega_a-\omega_b)$:

$$\mathbf{p}_{d-a+b}^{(3)}=\sum_{\gamma,\beta,\gamma',\alpha} i\chi_{d-a+b}\mathbf{p}_{3\gamma,2\beta}V_{2\beta,3\gamma'}^{(b)}V_{3\gamma'1\alpha}^{(a)*}V_{1\alpha,3\gamma}^{(d)}e^{i(\omega_d-\omega_a+\omega_b)t}+\text{c.c.}, \tag{4.2a}$$

where

$$\chi_{d-a+b}=\frac{T_{12}T_{23}}{\hbar^3[1-i(\omega_a-\omega_b-\omega_{21})T_{12}][1+i(\omega_d-\omega_a+\omega_b-\omega_{32})T_{23}]}\left[\frac{\Delta\tilde{\rho}_{32}T_{23}}{1+i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\tilde{\rho}_{31}T_{13}}{1-i(\omega_a-\omega_{31})T_{13}}\right], \tag{4.2b}$$

and the induced dipole moment per molecule at $\omega_c+(\omega_a-\omega_b)$:

$$\mathbf{p}_{c+a-b}^{(3)}=\sum_{\gamma,\alpha,\gamma',\beta} i\chi_{c+a-b}\mathbf{p}_{3\gamma,1\alpha}V_{1\alpha,3\gamma'}^{(a)}V_{3\gamma',2\beta}^{(b)*}V_{2\beta,3\gamma}^{(c)}e^{i(\omega_c+\omega_a-\omega_b)t}+\text{c.c.}, \tag{4.3a}$$

where

$$\chi_{c+a-b}=\frac{T_{12}T_{13}}{\hbar^3[1+i(\omega_a-\omega_b-\omega_{21})T_{12}][1+i(\omega_c+\omega_a-\omega_b-\omega_{31})T_{13}]}\left[\frac{\Delta\tilde{\rho}_{32}T_{23}}{1-i(\omega_b-\omega_{32})T_{23}}+\frac{\Delta\tilde{\rho}_{31}T_{13}}{1+i(\omega_a-\omega_{31})T_{13}}\right]. \tag{4.3b}$$

Let the primary light be the maser light at ω_l; we can study all the higher order Stokes and anti-Stokes radiations with the help of (4.2a), (4.2b), (4.3a), and (4.3b).

1. First Stokes Radiation

The following combination of frequencies can lead to the first Stokes radiation in the higher order Raman process involving four photons:

$$\omega_a=\omega_{AS},\quad \omega_b=\omega_l,\quad \omega_d=\omega_l,$$
$$\text{and} \tag{4.4}$$
$$\omega_c=\omega_d-\omega_a+\omega_b=\omega_l-\omega_{AS}+\omega_l=\omega_{S_1}.$$

Substituting (4.4) into (4.2a) and (4.2b), we obtain the induced dipole moment per molecule at ω_{S_1} produced in the process:

$$\mathbf{p}_{S_1}{}'^{(3)}=\sum_{\gamma,\beta,\gamma',\alpha} i\chi_{S_1}{}'\mathbf{p}_{3\gamma,2\beta}V_{2\beta,3\gamma'}^{(l)}V_{3\gamma',1\alpha}^{(AS)*}\times V_{1\alpha,3\gamma}^{(l)}e^{i\omega_{S_1}t}+\text{c.c.}, \tag{4.5a}$$

where

$$\chi_{S_1}{}'=\frac{T_{12}T_{23}}{\hbar^3[1-i(\omega_{AS}-\omega_l-\omega_{21})T_{12}][1+i(\omega_{S_1}-\omega_{32})T_{23}]}\times\left[\frac{\Delta\tilde{\rho}_{32}T_{23}}{1+i(\omega_l-\omega_{32})T_{23}}+\frac{\Delta\tilde{\rho}_{31}T_{13}}{1-i(\omega_{AS}-\omega_{31})T_{13}}\right] \tag{4.5b}$$

and use is made of the fact that $\omega_{AS}-\omega_l=\omega_l-\omega_{S_1}$.

Following exactly the same steps as used in the previous section, we find the average absorbed power per molecule:

$$P_{S_1}{}'=\text{Re}\{\sum_{\gamma,\beta,\gamma',\alpha} 2\omega_{S_1}\chi_{S_1}{}'V_{3\gamma,2\beta}^{(S_1)*}V_{2\beta,3\gamma'}^{(l)}\times V_{3\gamma',1\alpha}^{(AS)*}V_{1\alpha,3\gamma}^{(l)}\}. \tag{4.6}$$

By introducing the plane-wave spatial dependence of the fields, (3.7) and

$$V_{1\alpha,3\gamma'}^{(AS)}=\bar{V}_{1\alpha,3\gamma'}^{(AS)}\exp(-i\mathbf{k}_{AS}\cdot\mathbf{r}) \tag{4.7}$$

into (4.6) and then averaging over $\mathbf{r}$, we obtain the

[10] F. Bitter, Phys. Rev. 76, 833 (1949).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

average absorbed power per unit volume:

$$\langle P_{S_1}'\rangle_r = \mathrm{Re}\{\sum_{\gamma,\beta,\gamma',\alpha} 2N\omega_{S_1}\chi_{S_1}'\bar{V}_{3\gamma,2\beta}^{(S_1)*}\bar{V}_{2\beta,3\gamma'}^{(l)} \times \bar{V}_{3\gamma',1\alpha}^{(AS)*}\bar{V}_{1\alpha,3\gamma}^{(l)}\}\delta_{k_l+k_{l'}-k_{S1}-k_{AS}}. \quad (4.8)$$

In contrast to the ordinary Raman processes, in the present case there is a phase matching condition[4]:

$$\mathbf{k}_l+\mathbf{k}_{l'}-\mathbf{k}_{S_1}-\mathbf{k}_{AS}=0 \quad (4.8a)$$

in order for the higher order coherent Raman process to take place. Whether (4.8) corresponds to absorption or emission will depend upon the detailed nature of the molecular states and the relative phases of the fields involved, as expressed in the bracketed factor in (4.8). We shall examine some special cases later.

2. *First Anti-Stokes Radiation*

For the first anti-Stokes radiation we take the following combination of frequencies in (4.3a) and (4.3b):

$$\omega_a=\omega_l,\quad \omega_b=\omega_{S_1},\quad \omega_c=\omega_l,$$
$$\text{and} \quad \omega_d=\omega_l+\omega_l-\omega_{S_1}=\omega_{AS}. \quad (4.9)$$

The induced dipole moment per molecule at ω_{AS} produced in the higher order Raman process is

$$\mathbf{p}_{AS'}^{(3)} = \sum_{\gamma,\alpha,\gamma',\beta} i\chi_{AS}'\mathbf{p}_{3\gamma,1\alpha}V_{1\alpha,3\gamma'}^{(l)}V_{3\gamma',2\beta}^{(S_1)*} \times V_{2\beta,3\gamma}^{(l)}e^{i\omega_{AS}t}+\text{c.c.}, \quad (4.10a)$$

where

$$\chi_{AS}' = \frac{T_{12}T_{13}}{\hbar^3[1+i(\omega_l-\omega_{S_1}-\omega_{21})T_{12}][1+i(\omega_{AS}-\omega_{31})T_{13}]} \times\left[\frac{\Delta\bar{\rho}_{32}T_{23}}{1-i(\omega_{S_1}-\omega_{32})T_{23}}+\frac{\Delta\bar{\rho}_{31}T_{13}}{1+i(\omega-\omega_{31})T_{13}}\right] \quad (4.10b)$$

from (4.9) and (4.3a) and (4.3b). The average absorbed power per molecule is then

$$P_{AS}' = \mathrm{Re}\{\sum_{\gamma,\alpha,\gamma',\beta} 2\omega_{AS}\chi_{AS}'V_{3\gamma,1\alpha}^{(AS)*}V_{1\alpha,3\gamma'}^{(l)} \times V_{3\gamma',2\beta}^{(S_1)*}V_{2\beta,3\gamma}^{(l)}\}. \quad (4.11)$$

Upon introducing the spatial dependence of the fields, (3.7) and (4.7), and averaging over $\mathbf{r}$, we obtain the absorbed power per unit volume:

$$\langle P_{AS}'\rangle_r = \mathrm{Re}\{\sum_{\gamma,\alpha,\gamma',\beta} 2N\omega_{AS}\chi_{AS}'\bar{V}_{3\gamma,1\alpha}^{(AS)*}\bar{V}_{1\alpha,3\gamma'}^{(l)} \times \bar{V}_{3\gamma',2\beta}^{(S_1)*}\bar{V}_{2\beta,3\gamma}^{(l)}\}\delta_{k_l+k_{l'}-k_{S1}-k_{AS}}. \quad (4.12)$$

Again, one has the phase matching condition, (4.8a).[4]

3. *Higher Order Stokes and Anti-Stokes Radiations*

By taking suitable combinations of frequencies, properties of successive higher orders of Stokes and anti-Stokes radiations can be studied in much the same way using (4.2a), (4.2b), (4.3a), and (4.3b).

For example, for the second Stokes radiation, one takes the following independent combinations of frequencies:

$$\omega_a=\omega_l,\quad \omega_b=\omega_{S_1},\quad \omega_d=\omega_{S_1},$$
$$\omega_c=\omega_d-\omega_a+\omega_b=\omega_{S_1}-(\omega_l-\omega_{S_1})=\omega_{S_2}; \quad (4.13a)$$

and

$$\omega_a=\omega_{AS},\quad \omega_b=\omega_l,\quad \omega_d=\omega_{S_1},$$
$$\omega_c=\omega_{S_1}-(\omega_{AS}-\omega_l)=\omega_{S_2}. \quad (4.13b)$$

Physically, these correspond to the cases where a light wave at ω_{S_1} propagates in a medium which is modulated by the coherent molecular oscillations excited by two other light waves at ω_l and ω_{S_1}, (4.13a), or ω_{AS} and ω_l, (4.13b). As a result, absorption or emission at ω_{S_2} may take place. It is now obvious how the absorbed powers could be calculated. It may be of interest to look at the phase matching conditions.

It can be shown readily that for the first case, (4.13a), the phase matching condition is[4]

$$\mathbf{k}_{S_1}+\mathbf{k}_{S_1'}-\mathbf{k}_l-\mathbf{k}_{S_2}=0; \quad (4.14a)$$

and for the second case, (4.13b), it is

$$\mathbf{k}_{AS}-\mathbf{k}_l-\mathbf{k}_{S_1}+\mathbf{k}_{S_2}=0. \quad (4.14b)$$

It is clear that the general results given can be readily applied to a great variety of interesting experimental situations. We now study some specific simple cases.

B. Special Examples

Consider the first Stokes radiation in the higher order process. As in Sec. III, let us assume again that all the light waves are linearly polarized:

$$\bar{V}^{(l)}=p_z\bar{E}_l, \quad (4.15a)$$

$$\bar{V}^{(S_1)}=(p_x\sin\theta_{S_1}+p_z\cos\theta_{S_1})\bar{E}_{S_1}e^{i\psi_{S_1}}, \quad (4.15b)$$

$$\bar{V}^{(AS)}=(p_x\sin\theta_{AS}\cos\phi_{AS}+p_y\sin\theta_{AS}\sin\phi_{AS} +p_z\cos\theta_{AS})\bar{E}_{AS}e^{i\psi_{AS}}, \quad (4.15c)$$

in a suitably chosen coordinate system. With (4.15a)–(4.15c), we can evaluate the sums in (4.8), assuming that level 2 is a vibrational level; for example,

$$\Sigma=\sum_{\gamma,\beta,\gamma',\alpha}\bar{V}_{3\gamma,2\beta}^{(S_1)*}\bar{V}_{2\beta,3\gamma'}^{(l)}\bar{V}_{3\gamma',1\alpha}^{(AS)*}\bar{V}_{1\alpha,3\gamma}^{(l)}$$
$$=\sum_{\alpha,\beta}\bar{E}_l^2\bar{E}_{S_1}\bar{E}_{AS}e^{-i\psi_{S_1}-i\psi_{AS}}\langle 2,\beta|T(1,0)|3,\beta\rangle\langle 1,\alpha|T(1,0)|3,\alpha\rangle$$
$$\times[\langle 3,\alpha|T(1,0)|2,\beta\rangle\cos\theta_{S_1}+(1/\sqrt{2})(\langle 3,\alpha|T(1,-1)|2,\beta\rangle-\langle 3,\alpha|T(1,1)|2,\beta\rangle)\sin\theta_{S_1}]$$
$$\times[\langle 3,\beta|T(1,0)|1,\alpha\rangle\cos\theta_{AS}-(1/\sqrt{2})\langle 3,\beta|T(1,1)|1,\alpha\rangle(\sin\theta_{AS}\cos\phi_{AS}+i\sin\theta_{AS}\sin\phi_{AS})$$
$$+(1/\sqrt{2})\langle 3,\beta|T(1,-1)|1,\alpha\rangle(\sin\theta_{AS}\cos\phi_{AS}-i\sin\theta_{AS}\sin\phi_{AS})], \quad (4.16)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where the tensor operators are

$$T(1,0)=p_z,\quad T(1,-1)=\frac{1}{\sqrt{2}}p_x+\frac{i}{\sqrt{2}}p_y,\quad T(1,1)=-\frac{1}{\sqrt{2}}p_x+\frac{i}{\sqrt{2}}p_y. \tag{4.16a}$$

If, furthermore, levels 1 and 2 are nondegenerate, Eq. (4.16) simplifies to

$$\sum=\tfrac{1}{9}|p_{13}|^2|p_{23}|^2\bar{E}_l{}^2\bar{E}_{S_1}\bar{E}_{AS}e^{-i\psi_{S_1}-i\psi_{AS}}\cos\theta_{S_1}\cos\theta_{AS}. \tag{4.16b}$$

Assuming in addition that T_{13}, $T_{23}\to\infty$ and none of the single-photon frequencies are very near any of the transition frequencies, substituting (4.16b) into (4.8) and (4.5b), and taking the proper limits, we obtain for the average absorbed first Stokes radiation per unit volume:

$$\langle P_{S_1}{}'\rangle_r=-\frac{2N(\tilde{\rho}_{11}-\tilde{\rho}_{22})\omega_{S_1}T_{12}|p_{13}|^2|p_{23}|^2\bar{E}_l{}^2\bar{E}_{S_1}\bar{E}_{AS}}{9\hbar^3[1+(\omega_{AS}-\omega_l-\omega_{21})^2T_{12}{}^2]\Delta_1\Delta_2}\cos\theta_{S_1}\cos\theta_{AS}$$
$$\times[\cos(\psi_{S_1}+\psi_{AS})+T_{12}(\omega_{AS}-\omega_l-\omega_{21})\sin(\psi_{S_1}+\psi_{AS})]\delta_{k_l+k_{l'}-k_{Sl}-k_{AS}}, \tag{4.17}$$

where in the denominator, as before,

$$\Delta_1=\omega_{32}-\omega_{S_1}\quad\text{and}\quad\Delta_2=\omega_{32}-\omega_l\approx\omega_{31}-\omega_{AS}.$$

On resonance,[4] one has absorption of the first Stokes radiation, or $\langle P_{S_1}{}'\rangle_r\geq 0$, if $\pi/2<(\psi_{S_1}+\psi_{AS})<3\pi/2$, and emission of the first Stokes radiation, or $\langle P_{S_1}{}'\rangle_r\leq 0$, if $-\pi/2<(\psi_{S_1}+\psi_{AS})<\pi/2$.

Under similar circumstances, one finds the absorbed anti-Stokes radiation per unit volume:

$$\langle P_{AS}{}'\rangle_r=\frac{2N(\tilde{\rho}_{11}-\tilde{\rho}_{22})\omega_{AS}T_{12}|p_{13}|^2|p_{23}|^2\bar{E}_l{}^2\bar{E}_{S_1}\bar{E}_{AS}}{9\hbar^3[1+(\omega_{AS}-\omega_l-\omega_{21})^2T_{12}{}^2]\Delta_1\Delta_2}\cos\theta_{S_1}\cos\theta_{AS}$$
$$\times[\cos(\psi_{S_1}+\psi_{AS})-T_{12}(\omega_{AS}-\omega_l-\omega_{21})\sin(\psi_{S_1}+\psi_{AS})]\delta_{k_l+k_{l'}-k_{Sl}-k_{AS}}, \tag{4.18}$$

where, in the denominator, use is made of the fact that

$$\omega_l-\omega_{31}\approx\omega_{S_1}-\omega_{32}\quad\text{and}\quad\omega_{AS}-\omega_{31}\approx\omega_l-\omega_{32}.$$

Therefore, on resonance, when $\pi/2<(\psi_{S_1}+\psi_{AS})<3\pi/2$, $\langle P_{AS}{}'\rangle_r\leq 0$, and one has emission of anti-Stokes radiation; when $-\pi/2<(\psi_{S_1}+\psi_{AS})<\pi/2$, $\langle P_{AS}{}'\rangle_r\geq 0$ and one has absorption of anti-Stokes radiation.[4]

It follows also from (4.17) and (4.18) that on resonance and when the phase matching condition is satisfied

$$\langle P_{AS}{}'\rangle_r\cdot\omega_{AS}{}^{-1}=-\langle P_{S_1}{}'\rangle_r\cdot\omega_{S_1}{}^{-1}, \tag{4.19}$$

which is similar to the so-called Manley-Row relation, well-known in the theory of parametric amplifiers in microwave tube work and derived classically.[11]

One should note also that when, for example, three photons are simultaneously involved in the higher order Raman process, the net absorption of the first Stokes radiation is not just $\langle P_{S_1}{}'\rangle_r$. It should be $\langle P_{S_1}{}'\rangle_r$ plus $\langle P_{S_1}\rangle_r$, since both terms could be of the same order of magnitude. Similarly, the net absorption of anti-Stokes radiation is the sum of $\langle P_{AS}{}'\rangle_r$ and $\langle P_{AS}\rangle_r$. An examination of these expressions shows that when $\bar{E}_{S_1}>\bar{E}_{AS}$ one could have coherent amplification of anti-Stokes radiation of the right phase.

It is clear that specific results for other higher order Stokes and anti-Stokes radiations can be obtained in much the same way.

Most of the results and conclusions obtained in these simple examples, as well as those given in the previous section, have been given and discussed in detail in Ref. 4 on the basis of a phenomenological theory. In more complicated cases, the general results given in Sec. IIIA, as well as those given in Sec. II, will have to be used.

ACKNOWLEDGMENTS

The author would like to thank Dr. S. Kern for an interesting discussion on his experimental work and Dr. H. Statz for helpful discussions.

[11] P. K. Tien and H. Suhl, Proc. IRE 46, 700–706 (1958), with references to the original work of Manley and Row.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Interpretation of Frequency Shifts Due to Electron Exchange Collisions

P. L. BENDER
Joint Institute for Laboratory Astrophysics, Boulder, Colorado

(Received 6 January 1964)

Experiments by Hobart and Franken in 1961 demonstrated a shift in the free-electron spin-resonance frequency due to collisions with sodium atoms whose electron spins were polarized. The results have been interpreted on the assumptions that the S-wave phase shifts were proportional to the wave number k and that higher partial waves were not important. The triplet and singlet electron-sodium scattering lengths of Salmona and Seaton were used, but no check on the calculation was made since the sodium polarization was not known. Recently Balling, Hanson, and Pipkin have measured the shift for collisions with highly polarized rubidium atoms. They have also derived complete formulas for the frequency shift and line broadening in terms of the phase shifts with all partial waves included. In the present paper, an alternate derivation of the frequency shift and line broadening formulas is given and the experimental results are discussed on the basis of the phase-shift information that is available for hydrogen, sodium, and cesium. The small size of the observed ratio of frequency shift to line broadening for rubidium appears somewhat surprising in view of the larger ratios calculated for other alkalis.

1. INTRODUCTION

A MEASUREMENT of the free-electron g factor has been reported by Dehmelt.[1] His method involves polarization of the electrons by exchange collisions with sodium atoms which in turn have been polarized by optical pumping.[2,3] An exchange broadening cross section of $260\pi a_0^2$ or larger at 400°K was also measured. Salmona and Seaton[4] discussed this result on the basis of their calculated S-wave scattering lengths and found that the zero-energy-exchange cross section is large because the triplet and singlet scattering lengths are of opposite sign and are both large.

Hobart and Franken have repeated the electron g-factor experiment using a more homogeneous magnetic field.[5,6] During the experiment, they discovered a shift in the electron-spin-resonance frequency with the sodium polarization and density. The sample temperature was about 410°K and the magnetic field about 3.7 G. The pulsed discharge which produced the electrons was short and uncorrelated with the modulation frequency used to observe the electron resonance. An argon buffer gas at pressures of between a few Torr and about 50 Torr was present, so the electrons were presumably thermal. The frequency shift was interpreted as due to phase shifts in the off-diagonal elements of the electron-spin density matrix introduced by the electron-sodium exchange collisions. Maximum shifts of about 100 cps were observed under conditions where the electron resonance broadening by electron-sodium exchange collisions was about 2000 cps (half-width at half-maximum).[7]

The expected ratio of frequency shift to line broadening for the conditions of the Hobart and Franken experiment has been estimated[8] using Salmona and Seaton's triplet and singlet S-wave scattering lengths. The formula obtained, including only the S-wave contribution to the shift and broadening, was

$$(d'/w')=-P_{Na}\langle(\sin 2\Delta_0)\rangle_{\rm av}/\langle(1-\cos 2\Delta_0)\rangle_{\rm av}. \qquad (1)$$

Here d' and w' are the frequency shift and half-width, P_{Na} is the sodium electron polarization, $\Delta_0=\delta_0(1)-\delta_0(0)$ is the difference of the triplet and singlet S-wave phase shifts, and the average is over the thermal distribution of electron energies. If the scattering length approximation for the S-wave phase shifts is adequate and higher partial waves are not important, the observed ratio of shift to broadening requires a sodium polarization of about 10%. However, since the sodium polarization was not measured, the observations do not provide a check on the correctness of Eq. (1) or of the assumptions which were made.

Recently Balling, Hanson, and Pipkin[9] have measured shifts in the free-electron spin-resonance frequency due to collisions with polarized rubidium atoms. In this work the rubidium polarization was 90% or higher and the electrons were believed to be well thermalized. They also derived complete theoretical formulas for the shift and broadening of the electron spin resonance due to electron-atom exchange collisions in terms of the differences between the triplet and singlet scattering phase shifts for all partial waves. Their interpretation of the observed ratio of shift to half-width of -0.05 ± 0.01 was that the S-wave phase-shift difference was near some multiple of $(\pi/2)$ for electrons of the average energy, with $(\pi/2)$ itself most likely, and that higher partial waves were probably not important.

In Sec. 2 of the present paper an alternate derivation of the frequency shift and line broadening formulas for electron-atom exchange collisions is given. This deri-

[1] H. G. Dehmelt, Phys. Rev. **109**, 381 (1958).
[2] A. Kastler, J. Phys. Radium **11**, 255 (1950).
[3] H. G. Dehmelt, Phys. Rev. **105**, 1487 (1957).
[4] A. Salmona and M. J. Seaton, Proc. Phys. Soc. (London) **77**, 617 (1961).
[5] J. Hobart and P. Franken, International Conference on Optical Pumping, Heidelberg University, April, 1962 (unpublished abstract).
[6] J. L. Hobart, thesis, University of Michigan, 1962 (unpublished).
[7] P. Franken (private communication).
[8] P. L. Bender, International Conference on Optical Pumping, Heidelberg University, April, 1962 (unpublished abstract).
[9] L. C. Balling, R. J. Hanson, and F. M. Pipkin, Phys. Rev. **133**, A607 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

vation is based on the expressions one would obtain from Baranger's general impact theory[10] if the triplet and singlet states were energy eigenstates during the time between collisions. The proof that the resulting formulas hold for arbitrary alkali nuclear spin is given in the Appendix.

In Sec. 3 the available theoretical information on phase shifts for electron-alkali scattering is reviewed. The modified effective range theory expansions of the S-wave phase shifts in terms of the wave number k appear useful for hydrogen and sodium but break down for cesium in the temperature range of interest. Phase-shift calculations as a function of energy for hydrogen and cesium and the scattering lengths for sodium are used in discussing the frequency-shift ratio for rubidium. The experimentally observed result is not easy to understand in this way, and thus more extensive phase-shift calculations as well as further frequency-shift experiments with cesium and sodium appear to be desirable.

2. ALTERNATE DERIVATION OF ELECTRON SPIN RESONANCE FREQUENCY SHIFT

For the experimental conditions of Hobart and Franken[5] and of Balling, Hanson, and Pipkin,[9] the alkali and buffer gas densities are low enough so that the impact approximation is very good. The number of electrons is much less than the number of alkali atoms, and we thus consider only electron-alkali collisions. Magnetic interaction energies are small enough to be neglected during collisions. We take the axis of quantization as along the magnetic-field direction. The components of the free-electron spin and the valence-electron spin along the axis are labeled as m_e and m_J, respectively, while the total electron spin and its component along the axis are S and m_S. We assume that the alkali nuclear spin is zero for simplicity and show in the Appendix that the result is independent of the nuclear spin. However, we do require that the off-diagonal elements of the alkali atom density matrix in the energy (i.e., F, M_F) representation between collisions be negligible.

If the triplet and singlet states of the coupled free electron and alkali-valence electron were nondegenerate energy eigenstates between collisions, we could use the general impact theory results of Baranger[10] to obtain the shift and broadening of the transition between these levels in terms of the scattering phase shifts $\delta_l(1)$ and $\delta_l(0)$ for the triplet and singlet states, respectively. The subscript refers to the lth partial wave. Taking the triplet as the higher level, the half-width and the shift are given, respectively, in angular frequency by

$$W=\left\langle \frac{1}{2}nv\int d\Omega|f_1(\Omega)-f_0(\Omega)|^2\right\rangle_{\rm av}, \tag{2}$$

$$d=\Big\langle -(2\pi nv/k)R[f_1(0)-f_0(0)]$$

$$+\tfrac{1}{2}nvi\int d\Omega[f_0^*(\Omega)f_1(\Omega)-f_0(\Omega)f_1^*(\Omega)]\Big\rangle_{\rm av}. \tag{3}$$

Here $f_1(\Omega)$ and $f_0(\Omega)$ are the triplet and singlet elastic scattering amplitudes as a function of direction, R means the real part, n is the number of scatterers per a_0^3, v is the electron velocity in a_0/sec, a_0 is the Bohr radius for hydrogen, and k^2 is the electron energy in units of 13.6 eV. The average is over the electron-velocity distribution, which is assumed to be Maxwellian. Since

$$f_S(\theta)=(2ik)^{-1}\sum_l(2l+1)(e^{2i\delta_l(S)}-1)P_l(\cos\theta), \tag{4}$$

where $S=1$ and 0 correspond to the triplet and singlet amplitudes, we find

$$w=\langle(\pi nv/k^2)\sum_l(2l+1)[1-\cos2\Delta_l]\rangle_{\rm av}, \tag{5}$$

$$d=\langle-(\pi nv/k^2)\sum_l(2l+1)[\sin2\Delta_l]\rangle_{\rm av}, \tag{6}$$

where $\Delta_l=\delta_l(1)-\delta_l(0)$. Converting n and v to cgs units and labeling the most probable electron velocity as v_m and the corresponding value of k as k_m, we have

$$w=\left(\frac{\pi a_0^2nv_m}{k_m^2}\right)\sum_l(2l+1)\langle(1-\cos2\Delta_l)/(v/v_m)\rangle_{\rm av}, \tag{7}$$

$$d=-\left(\frac{\pi a_0^2nv_m}{k_m^2}\right)\sum_l(2l+1)\langle(\sin2\Delta_l)/(v/v_m)\rangle_{\rm av}. \tag{8}$$

This is equivalent to the following collisional rates of change of the density matrix elements ${}^e\sigma_{ij}$ in the electron-spin-coupled representation:

$$(d{}^e\sigma_{ij}/dt)_{\rm Col.}=-(w-id){}^e\sigma_{ij} \tag{9}$$

for i corresponding to a singlet state and j to a triplet state, and

$$(d{}^e\sigma_{ij}/dt)_{\rm Col.}=0 \tag{10}$$

for i and j both corresponding to singlet or both to triplet states.

In reality, the states between which transitions are observed are the free-electron spin states. We will let ρ be the free-electron spin density matrix in the usual representation. Ignoring the off-diagonal elements for the alkali spin density matrix u and also ignoring correlations between the free-electron and alkali-electron density matrices, we can write the total spin

[10] M. Baranger, Phys. Rev. 111, 481 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

density matrix $\sigma=\rho\times u$ as

M_e, M_J	$\frac{1}{2}, \frac{1}{2}$	$-\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$	$-\frac{1}{2}, -\frac{1}{2}$
$\frac{1}{2}, \frac{1}{2}$	$A\rho_{11}$	$A\rho_{12}$	0	0
$-\frac{1}{2}, \frac{1}{2}$	$A\rho_{21}$	$A\rho_{22}$	0	0
$\frac{1}{2}, -\frac{1}{2}$	0	0	$B\rho_{11}$	$B\rho_{12}$
$-\frac{1}{2}, -\frac{1}{2}$	0	0	$B\rho_{21}$	$B\rho_{22}$

$\sigma=$. (11)

Transforming to the electron-spin-coupled representation by the transformation

$$^c\sigma=\Gamma\sigma\Gamma', \tag{12}$$

where

S, M_S \ M_e, M_J	$\frac{1}{2}, \frac{1}{2}$	$-\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$	$-\frac{1}{2}, -\frac{1}{2}$
1, 1	1			
1, 0		$1/\sqrt{2}$	$1/\sqrt{2}$	
1, −1				1
0, 0		$-1/\sqrt{2}$	$1/\sqrt{2}$	

$\Gamma=$ (13)

gives

S, M_S \ S, M_S	1, 1	1, 0	1, −1	0, 0
1, 1	⋯	⋯	⋯	$-(A\rho_{12})/\sqrt{2}$
1, 0	⋯	⋯	⋯	$(\rho_{11}-A)/2$
1, −1	⋯	⋯	⋯	$(B\rho_{21})/\sqrt{2}$
0, 0	$-(A\rho_{21})/\sqrt{2}$	$(\rho_{11}-A)/2$	$(B\rho_{12})/\sqrt{2}$	⋯

$^c\sigma=$. (14)

The other elements are not necessarily zero but are not needed. From Eq. (9), setting $p=\frac{1}{2}(w-id)$,

S, M_S \ S, M_S	1, 1	1, 0	1, −1	0, 0
1, 1	0	0	0	$\sqrt{2}p^*A\rho_{12}$
1, 0	0	0	0	$-p^*(\rho_{11}-A)$
1, −1	0	0	0	$-\sqrt{2}p^*B\rho_{21}$
0, 0	$\sqrt{2}pA\rho_{21}$	$-p(\rho_{11}-A)$	$-\sqrt{2}pB\rho_{12}$	0

$(d^c\sigma/dt)_{\text{Col.}}=$. (15)

Returning to the uncoupled representation by the transformation

$$(d\sigma/dt)_{\text{Col.}}=\Gamma'(d^c\sigma/dt)_{\text{Col.}}\Gamma \tag{16}$$

gives

M_e, M_J \ M_e, M_J	$\frac{1}{2}, \frac{1}{2}$	$-\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$	$-\frac{1}{2}, -\frac{1}{2}$
$\frac{1}{2}, \frac{1}{2}$	0	$-p^*A\rho_{12}$	$p^*A\rho_{12}$	0
$-\frac{1}{2}, \frac{1}{2}$	$-pA\rho_{21}$	$(w/2)(\rho_{11}-A)$	$-i(d/2)(\rho_{11}-A)$	$pB\rho_{12}$
$\frac{1}{2}, -\frac{1}{2}$	$pA\rho_{21}$	$i(d/2)(\rho_{11}-A)$	$-(w/2)(\rho_{11}-A)$	$-pB\rho_{12}$
$-\frac{1}{2}, -\frac{1}{2}$	0	$p^*B\rho_{21}$	$-p^*B\rho_{21}$	0

$(d\sigma/dt)_{\text{Col.}}=$. (17)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Finally, by contracting we obtain

$$(d\rho/dt)_{\text{Col.}}=\begin{array}{c|cc} M_e \backslash M_e & \frac{1}{2} & -\frac{1}{2} \\ \hline \frac{1}{2} & -\frac{1}{2}w(\rho_{11}-A) & -[\frac{1}{2}w+\frac{1}{2}id(A-B)]\rho_{12} \\ -\frac{1}{2} & -[\frac{1}{2}w-\frac{1}{2}id(A-B)]\rho_{21} & \frac{1}{2}w(\rho_{11}-A) \end{array}. \tag{18}$$

The above collisional rates of change for the off-diagonal elements of ρ correspond to the following half-width and frequency shift for the free-electron spin resonance:

$$w'=\tfrac{1}{2}w, \tag{19}$$

$$d'=\tfrac{1}{2}P_A d, \tag{20}$$

where $P_A=A-B$ is the alkali electron spin polarization. These formulas are the same as those obtained by Balling, Hanson, and Pipkin[9] by essentially repeating Baranger's calculation for the case of electron exchange collisions. They have been rederived here mainly for pedagogic reasons. The ratio of shift to broadening for the case where only S-wave collisions are important agrees with our earlier estimated value except that the velocity averaging procedure is now specified.

3. LOW-ENERGY ELECTRON-ALKALI ELASTIC SCATTERING

We are now ready to review the available information on low-energy electron-alkali elastic scattering phase shifts. The only calculations available which extend over the necessary energy range are for hydrogen and cesium. For sodium, only the S-wave scattering lengths are available.

We define

$${}^sI_l\equiv\langle(\sin 2\Delta_l)/(v/v_m)\rangle_{\text{av}}, \tag{21}$$

$${}^cI_l\equiv\langle(1-\cos 2\Delta_l)/(v/v_m)\rangle_{\text{av}}, \tag{22}$$

$$R\equiv(-d'/P_A w'), \tag{23}$$

$$x\equiv(v/v_m)=(k/k_m). \tag{24}$$

Then,

$$R=[\sum_l(2l+1){}^sI_l]/[\sum_l(2l+1){}^cI_l], \tag{25}$$

where

$${}^sI_l=(4/\sqrt{\pi})\int_0^\infty x[\sin 2\Delta_l(k)]e^{-x^2}dx, \tag{26}$$

$${}^cI_l=(4/\sqrt{\pi})\int_0^\infty x[1-\cos 2\Delta_l(k)]e^{-x^2}dx. \tag{27}$$

For hydrogen and cesium, we will use $T=290°\text{K}$, corresponding to the rubidium experiments of Balling, Hanson, and Pipkin,[9] while for sodium we will also use $T=410°\text{K}$, corresponding roughly to the sodium experiments of Hobart and Franken[5,6] and of Dehmelt.[1] We have $k_m(290°\text{K})=0.0429$ and $k_m(410°\text{K})=0.0510$.

For the hydrogen S wave, we take the results of Schwartz.[11] However, his lowest points are for $k^2=0.01$ and 0.02. We must therefore use an interpolation formula. O'Malley, Spruch, and Rosenberg[12,13] have shown that the normal effective range theory expansion is not valid in the presence of a long-range potential such as the polarization potential $V(r)=-\alpha/r^4$. However, a modified effective range theory expansion which is valid for atoms which have spherically symmetric ground states has been given by these authors. From Eq. (2.3) of O'Malley,[14] we obtain

$$\Delta_0=[A(1)-A(0)]k+\tfrac{4}{3}[A(1)-A(0)](\alpha/a_0^3)k^3\ln k+0(k^3), \tag{28}$$

$$\Delta_1=0(k^3). \tag{29}$$

Here $A(1)$ and $A(0)$ are the triplet and singlet S-wave scattering lengths and α is the atomic polarizability. From Schwartz's paper $[A(1)-A(0)]=4.196$. Using the hydrogen polarizability[15] of $4.5a_0^3$ to calculate the coefficient of $k^3\ln k$ and including a k^3 term, we obtain

$$\Delta_0=4.196k+25.2k^3\ln k+24k^3, \tag{30}$$

where the coefficient of k^3 has been fit to Schwartz's value at $k=0.1$. With this interpolation formula, ${}^sI_0=0.339$ and ${}^cI_0=0.066$.

For the P wave we use the results of Burke and Schey.[16] Including only a k^3 term and fitting to their lowest point, we find $\Delta_1=3.8k^3$, ${}^sI_1=0.001$, ${}^cI_1=0.000$. Also from Burke and Schey, higher partial waves are negligible. Thus, $R_H(290°\text{K})=5.2$.

For cesium, the only calculated phase shifts which appear to be useful for the present problem are the very recent ones of Stone and Reitz.[17] Values of Δ_0, Δ_1, and Δ_2 obtained from their results are given in Table I. It appears that the first three or four terms of the modified effective range theory expansion are not adequate to represent the results even at moderate values of k. This is not surprising in view of the high

[11] C. Schwartz, Phys. Rev. 124, 1468 (1961).
[12] T. F. O'Malley, L. Spruch, and L. Rosenberg, J. Math. Phys. 2, 491 (1961).
[13] T. F. O'Malley, L. Rosenberg, and L. Spruch, Phys. Rev. 125, 1300 (1962).
[14] T. F. O'Malley, Phys. Rev. 130, 1020 (1963).
[15] N. F. Mott and I. N. Sneddon, *Wave Mechanics and its Applications* (Clarendon Press, Oxford, 1948).
[16] P. G. Burke and H. M. Schey, Phys. Rev. 126, 147 (1962).
[17] P. M. Stone and J. R. Reitz, Phys. Rev. 131, 2101 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Triplet-singlet phase-shift differences Δ_l for the S, P, and D partial-wave scattering of electrons from cesium.[a] The functions $f_0=(0.78)[1-\exp(-36k)]$ and $f_1=-(0.66k)\times[1-\exp(-300k^2)]$ are chosen to approximate Δ_0 and Δ_1 in the energy range of interest.

k^2	Δ_0	Δ_1	Δ_2	f_0	f_1
0.001	0.55	−0.05	0.00	0.53	−0.05
0.003	0.65	−0.21	0.01	0.67	−0.21
0.005	0.70	−0.37	0.01	0.72	−0.36
0.007	0.74	−0.49	0.00	0.74	−0.48
0.010	0.77	−0.64	0.03	0.76	−0.63
0.013	0.79	−0.75	0.04	0.77	−0.74
0.020	0.83	−0.93	0.15	0.77	−0.93

[a] See Ref. 17.

polarizability of cesium.[18] Instead, we represent Δ_0 and Δ_1 by functions $f_0=(0.78)[1-\exp(-36k)]$ and $f_1=-(0.66)k[1-\exp(-300k^2)]$ which give adequate agreement with Δ_0 and Δ_1 over the range of interest and which have leading terms proportional to k and k^3, respectively, near zero energy. These leading terms agree with the modified effective range theory. If one interprets the coefficient of the leading term in f_0 as a rough approximation to the triplet-singlet scattering-length difference, this gives $\Lambda(1)-\Lambda(0)=28$. Using f_0 and f_1, we find for cesium ${}^3I_0=0.94$, ${}^cI_0=0.61$, ${}^3I_1=-0.03$, and all other contributions are 0.01 or less. Thus, we obtain $R_{Cs}(290°K)\simeq 1.4$. Since the contributions from P and higher partial waves are small for cesium as well as hydrogen, the S-wave approximation is presumably good for the other alkalis also.

A possible difficulty with the interpretation given above appears when one looks at the calculated triplet S-wave phase shifts rather than just the phase-shift differences. The triplet phase shifts decrease fairly slowly, but the first one is 0.37 rad above 3π. Since the zero-energy limit must be a multiple of π, this raises the possibility that the zero-energy value consistent with the finite energy values is 4π. If so, all the values of Δ_0 in Table I should be increased by π. This would change only the way in which one extrapolates from the lowest calculated point to zero energy and would probably not change the value of R by more than 20%.

Actually, a scattering length difference large enough to give such a steep initial slope for Δ_0 as a function of k seems questionable. Some support for this view comes from the fact that the P-wave triplet phase shifts are also slowly decreasing and the lowest energy one is 0.05 rad above 2π. Since a zero-energy limit of anything but 2π seems inconceivable, it appears that the low-energy values for the S-wave phase shifts may also be somewhat offset by the computational procedure used by Stone and Reitz. This makes the uncertainty in Δ_0 appear to be of the order of 0.4 rad or larger and raises the possibility that Δ_0 could approach or go through $(\pi/2)$ at thermal energies and give a very small value for R_{Cs}.

[18] G. E. Chamberlain and J. C. Zorn, Phys. Rev. 129, 677 (1963).

In the case of sodium the main pieces of information available are the triplet and singlet S-wave scattering lengths $\Lambda(1)=9$ and $\Lambda(0)=-12$ calculated by Salmona and Seaton. If the scattering length approximation under which $\delta_0(1)=k\Lambda(1)$, $\delta_0(0)=k\Lambda(0)$ were valid, we would have $\Delta_0=21k$. Including only the S-wave contribution, this gives ${}^3I_0=0.68$, ${}^cI_0=1.28$, $R_{Na}=0.53$, at 410°K while at 290°K ${}^3I_0=0.80$, ${}^cI_0=1.10$, $R_{Na}=0.73$. In the same approximation, defining

$$\bar{Q}_{ex}=\langle(vQ_{ex}(v))\rangle_{av}/(\bar{v}), \tag{31}$$

we find,

$$\bar{Q}_{ex}=\pi a_0^2(\sqrt{\pi}\,{}^cI_0/4k_m^2). \tag{32}$$

At 410°K this gives $\bar{Q}_{ex}=220\pi a_0^2$. Dehmelt's experimental result,[1] adjusted to correspond to $\bar{v}$ instead of v_{rms}, is $Q_{ex}\geq 280\pi a_0^2$. Salmona and Seaton's zero energy cross section is $Q_{ex}=440\pi a_0^2$.

We will next calculate the modified effective range theory expansion for sodium. From Eq. (4.9) of O'Malley, Rosenberg, and Spruch[13] we find

$$\Delta_0=[\Lambda(1)-\Lambda(0)]k+Ak^3\ln k+Bk^3+Ck^4+\cdots. \tag{33}$$

Here,

$$A=\tfrac{4}{3}(\alpha/a_0^3)[\Lambda(1)-\Lambda(0)], \tag{34}$$

$$B=\{A\cdot\ln[1.23(\alpha/a_0^3)^{1/2}]+\tfrac{1}{3}\pi(\alpha/a_0^3)^{1/2}[\Lambda^2(0)-\Lambda^2(1)]-\tfrac{1}{3}[\Lambda^3(1)-\Lambda^3(0)]+\tfrac{1}{2}[\Lambda^2(0)r_{poS}-\Lambda^2(1)r_{poT}]/a_0\}, \tag{35}$$

$$C=\tfrac{2}{3}\pi(\alpha/a_0^3)[\Lambda^2(1)-\Lambda^2(0)], \tag{36}$$

where r_{poT} and r_{poS} are the triplet and singlet modified effective ranges and α is the polarizability. Using[18] $\alpha=21\times10^{-24}$ cm^3, we calculate $A=3970$, $C=-18\,700$, and $B=10\,600+72(r_{poS}/a_0)-40.5(r_{poT}/a_0)$. Transforming to the variable $x=(k/k_m)$

$$\Delta_0(410°K)=1.07x-0.16x^3+[0.0095(r_{poS}/a_0)-0.0054(r_{poT}/a_0)]x^3+0.52x^3\ln x-0.127x^4+\cdots. \tag{37}$$

We do not know the modified effective ranges, but at least one of them must be very large in order to make the coefficient of x^3 approach one. Since the other coefficients of x^3 and higher terms are fairly small, it appears that the modified range theory expansion may be useful for sodium at 410°K. The exchange cross section and the ratio of electron frequency shift to line broadening calculated in the scattering length approximation thus may be fairly accurate.

Another procedure is to approximate the sodium phase-shift differences by the same function of (k/k_{ex}) as was used to fit the cesium results of Stone and Reitz. Here k_{ex}^2 is the excitation energy for the atom in question. The reason for trying this is that the values of Δ_0 for cesium derived from Stone and Reitz are rather similar to those for hydrogen if both are plotted as a function of (k/k_{ex}). At 410°K we obtain ${}^cI_0\simeq 0.94$, ${}^cI_0\simeq 0.60$, $R_{Na}\simeq 1.6$, and $\bar{Q}_{ex}\simeq 100\pi a_0^2$. This procedure

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

gives poor agreement with Dehmelt's quoted lower limit for the experimental exchange cross section. While Dehmelt's number is not actually a lower limit, since no uncertainties are included, it would indicate a definite error in the experiment if the cross section were as low as $100\pi a_0^2$. Thus, the somewhat similar behavior of the presently available phase-shift differences for cesium and hydrogen does not appear to be characteristic of sodium.

One other possible source of information concerning the phase-shift difference for sodium comes from a rough estimate of the sodium polarization in Hobart's experiment based on a trace he gives[6] of the sodium Zeeman resonances at 3.68 G. This trace (Fig. 3.10 of Ref. 6) indicates a ratio of roughly 2.8:1 for the strongest and weakest $F=2$ Zeeman components. Assuming that the sublevel populations are independent of F and that sublevels differing in M_F by 1 have a constant ratio of populations leads to an approximate sodium-electron polarization $P_A=-0.17$. These assumptions correspond to spin-exchange equilibrium between the sodium atoms. Actually, the small amplitude of the $F=1$ resonances indicates that the magnitude of the polarization was probably higher. However, since we do not know precisely what correction to make, we will use the above number. Combining this with the value $R_{\mathrm{Na}}=0.53$ calculated on the scattering length approximation using the scattering lengths of Salmona and Seaton gives a ratio of shift to half-width of 0.09. While the maximum ratio of shift to half-width observed was only 0.05, it is not clear that the trace given by Hobart applies to the same experimental conditions. Thus, little can be said about the agreement of theory and experiment for sodium. However, it would not be surprising if the sodium phase-shift difference curve actually rose somewhat more rapidly than the scattering length approximation curve which we have used.

4. DISCUSSION

We are now ready to discuss the results of Balling, Hanson, and Pipkin.[9] Their experimental value of $R=0.05\pm0.01$ for rubidium at about 290°K is more than an order of magnitude less than the corresponding value for cesium calculated from the phase shifts of Stone and Reitz.[17] If the phase-shift difference curve for rubidium had the same shape as that calculated for cesium, it would have to have more than twice as high a maximum in order to agree with experiment. However, it seems uncertain whether a calculation for rubidium done by a method similar to that of Stone and Reitz for cesium would give this much different a result. Also, it appears doubtful whether the rather high value for the cesium polarizability used by Stone and Reitz can account for the possible discrepancy.

The argument given by Balling, Hanson, and Pipkin that the Rb^- ion is just barely bound and that therefore $\delta_0(0)$ should be about $\pi/2$ appears to be open to question. This argument, which was also used by Dehmelt[1] for Na, is based on the result for a deep square-well potential[19] and is only valid if the binding energy is small compared with the incident electron energy. The modified effective range theory expansion[12,13] shows that the square well is not a good approximation at extremely low energies. Also, the binding energies of alkali negative ions are believed to be much higher than thermal energies.[20] In particular, a calculation by Szasz[21] gives binding energies of 0.43, 0.53, 0.49, 0.43, and 0.42 eV for Li, Na, K, Rb, and Cs. These results are not expected to be very accurate, as indicated by the value for Li of 0.62 eV obtained by Weiss in a much more thorough calculation.[20] However, the relative values may be good and fairly strong binding for Rb^- seems likely.

In view of the possible discrepancy mentioned above, the assumptions made in the frequency-shift calculation were reviewed. The least certain ones appeared to be that the alkali matrix in the F, M_F representation can be taken as diagonal and that direct interactions between the free electrons can be neglected. The first assumption is discussed briefly at the end of the Appendix and appears to hold for the rubidium experiments. The second assumption is more difficult to investigate because of the very long range of the electron-electron interaction. However, if such interactions were important one might expect that normal variations of the experimental conditions would have given erratic results.

Assuming that the experimental results and their interpretation in terms of the scattering phase shifts are correct, we are left with two possible hypotheses. One is that the rubidium S-wave phase-shift difference Δ_0 increases rapidly at very low energies and fortuitously passes through or approaches $(\pi/2)$ at the proper energy to make R small. This seems somewhat unlikely but cannot be ruled out. The other possibility is that Δ_0 for the higher alkalis rises rapidly initially but for some reason flattens out near $(\pi/2)$ at roughly thermal energies. This would require that the low-energy phase-shift differences for cesium taken from Stone and Reitz be off by approximately 1 rad.

5. CONCLUSIONS

In view of the large difference between the experimental ratio of shift to broadening for rubidium and the ratio calculated for cesium using the phase shifts of Stone and Reitz, the present situation appears to be somewhat unsatisfactory. The quickest approach to seeing if there is a real discrepancy between experiment and theory would be an experimental measurement on cesium.[22] However, in view of the difficulty of calcu-

[19] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon Press, Oxford, 1949), 2nd ed.

[20] L. M. Branscomb, in *Atomic and Molecular Processes*, edited by D. R. Bates (Academic Press Inc., New York, 1962).

[21] L. Szasz, Acta Phys. Acad. Sci. Hung. 6, 307 (1956).

[22] *Note added in proof.* Recent results for the ratio of frequency shift to line broadening in cesium appear to be similar to those for rubidium: L. C. Balling and F. M. Pipkin, Bull. Am. Phys. Soc. 9, 91 (1964). There is thus a definite discrepancy with the cesium phase-shift calculation of Stone and Reitz.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

lations on cesium, it also appears to be desirable to repeat the experiments for sodium or for hydrogen.

ACKNOWLEDGMENTS

I would like to acknowledge helpful conversations with T. F. O'Malley on low-energy electron-alkali scattering phase shifts. Thanks are also due to P. A. Franken, P. M. Stone, and F. M. Pipkin for communicating their results prior to publication.

APPENDIX: DERIVATION OF ELECTRON SPIN RESONANCE FREQUENCY SHIFT FOR ARBITRARY NUCLEAR SPIN

We will treat the case of electrons with an arbitrary 2×2 spin-density matrix ρ colliding with alkali atoms having an electron spin of $\frac{1}{2}$ and a nuclear spin I. ρ is taken to be in the M_e representation, where M_e is the free-electron spin component along the axis and the two states are ordered according to decreasing M_e. The alkali spin density matrix u_{F,M_F} in the F, M_F representation is assumed to be diagonal. The alkali sublevels are taken to be listed in order of decreasing M_F and subordered according to decreasing F; i.e., the F, M_F order of listing is $[(I+\frac{1}{2}, I+\frac{1}{2}), (I+\frac{1}{2}, I-\frac{1}{2}), (I-\frac{1}{2}, I-\frac{1}{2}), (I+\frac{1}{2}, I-\frac{3}{2}), \cdots]$. When u_F is transformed to the M_I, M_J representation in which the states are ordered according to decreasing M_I and subordered according to decreasing M_J, the resulting matrix u will have off-diagonal elements connecting only the second and third, the fourth and fifth, etc., states. This is because there can be no off-diagonal elements connecting states with different values of (M_J+M_I). The diagonal elements can be labeled $(A_I, B_I, A_{I-1}, B_{I-1}, \cdots)$, where A corresponds to $M_J=\frac{1}{2}$, B to $M_J=-\frac{1}{2}$, and the subscript to M_I. Note that the expectation value of the electron polarization for the atom is given by

$$P_A=\sum_{M_I=-I}^{I}(A_{M_I}-B_{M_I}). \tag{38}$$

We now form the density matrix σ for the electron-plus-atom system in the (M_e, M_J, M_I) representation by taking the direct product of ρ and of u in the (M_J, M_I) representation. The states of σ are ordered according to decreasing M_I, first subordered according to decreasing M_J, and second subordered according to decreasing M_e. σ can be thought of as made up of $(2I+1)^2$ 4×4 submatrices. The 4×4 submatrices σ_{M_I} along the diagonal of σ which correspond to different M_I are of the form given in Eq. (11) except that A and B are replaced by A_{M_I} and B_{M_I}.

The transformation of σ to the representation with the electron spins coupled and the states ordered by decreasing M_I, first subordered by decreasing values of the total electron spin S, and second subordered by decreasing values of M_J is given by the following transformation:

$$^c\sigma=\Gamma_I\sigma\Gamma_I'. \tag{39}$$

Γ_I is made up of $2I+1$ identical 4×4 submatrices Γ along its diagonal, with Γ given by Eq. (13). We find then that the $2I+1$ 4×4 submatrices along the diagonal of σ^c are given by

$$^c\sigma_{M_I}=\Gamma\sigma_{M_I}\Gamma'. \tag{40}$$

We can now make use of the results for the case of zero nuclear spin. When we find $(d\sigma/dt)_{\text{Col.}}$ we will contract it to obtain $(d\rho/dt)_{\text{Col.}}$ as was done in Sec. 2. The only parts of $(d\sigma/dt)_{\text{Col.}}$ which we will need are the 4×4 submatrices $(d\sigma_{M_I}/dt)_{\text{Col.}}$. However, since

$$(d\sigma_{M_I}/dt)_{\text{Col.}}=\Gamma'(d^c\sigma_{M_I}/dt)_{\text{Col.}}\Gamma \tag{41}$$

and since $(d^c\sigma_{M_I}/dt)_{\text{Col.}}$ is obtained from $^c\sigma_{M_I}$ in the same way that $(d^c\sigma/dt)_{\text{Col.}}$ was obtained from $^c\sigma$ in Sec. 2, the contribution to $(d\rho/dt)_{\text{Col.}}$ from each σ_{M_I} is identical to that given in Eq. (18) except with A and B replaced by A_{M_I} and B_{M_I}. Thus, $(d\rho/dt)_{\text{Col.}}$ is the same as for the case of zero nuclear spin except that $A\rightarrow\sum_{M_I}A_{M_I}$ and $B\rightarrow\sum_{M_I}B_{M_I}$. In particular, $w'=\frac{1}{2}w$ and $d'=\frac{1}{2}dP_A$, which are identical to Eqs. (19) and (20) for the case of zero nuclear spin.

In this calculation it was assumed that u_{F,M_F} was diagonal. However, u_{F,M_F} does have two off-diagonal elements which have approximately the same time dependence as the off-diagonal elements of ρ. While the rf field at the electron spin resonance frequency cannot directly produce finite values for these elements because the corresponding transition would require $\Delta M_I=2I+1$, these elements can still be produced by the electron-alkali collisions. Thus, in order to be sure that off-diagonal elements of u_{F,M_F} do not change the electron frequency shift and line broadening, we must show that these elements are much smaller than the off-diagonal elements of ρ.

A direct calculation for arbitrary nuclear spin appears to be difficult. However, the calculation can be carried through for zero nuclear spin with the special condition that the rf field does not affect the alkali atom directly. In this case, it is found that the condition for the off-diagonal elements of u to be much smaller than those of ρ is as follows: The ratio of alkali density to electron density must be large compared with the ratio of the width of the electron spin resonance due to alkali collisions compared with the reciprocal of the decay time for the alkali off-diagonal elements. If the decay times for the off-diagonal elements of both ρ and u were determined by electron-alkali exchange collisions, this condition would not hold. However, as long as the decay time for the alkali off-diagonal elements is determined mainly by some other cause such as light absorption rather than by electron-alkali exchange collisions, then the assumption that u is diagonal should be quite good. This condition was apparently satisfied in the experiments of Balling, Hanson, and Pipkin, and of Hobart and Franken.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Fine Structure of Helium*†

CHARLES SCHWARTZ
Department of Physics, University of California, Berkeley, California
(Received 6 January 1964)

Under the expectation that experiments will soon give values for the fine structure intervals of the 2^3P state of helium to an accuracy of $1/10^6$, we have undertaken a program of calculations which, it is hoped, will lead to a new determination of the fine-structure constant α. This paper gives a brief survey of the over-all program, and a detailed report of the successful completion of the first task: The construction of approximate solutions to the Schrödinger equation which lead to average values of the leading fine structure operators accurate to about one part per million.

I. INTRODUCTION

THE early studies of the fine structure of the helium atom by Breit[1] provided very important confirmation of the newly developing theory of the quantum interactions of electrons and the electromagnetic field. Over the intervening years there have been several efforts to increase the accuracy of this analysis, chiefly by the construction of successively more complicated approximations to the nonrelativistic 2-electron wave function. Breit,[1] with a 2-term trial function got values of the fine-structure intervals ν (see Fig. 1) accurate to several percent. Araki[2] and co-workers with 8 terms came within about 1%, and also considered two small corrections to the first-order theory. Traub and Foley,[3] with an 18-parameter function reduced the error to about one part in 10^3; and most recently Pekeris, Schiff, and Lifson,[4] with up to 220 terms in the trial function, came within one part in 10^4 of the final results we have attained. (These numbers refer to the larger interval ν_{01}. In the calculation of the small interval ν_{12} there is a cancellation between separate terms amounting to a full order of magnitude, and a corresponding loss of accuracy results.)

While these studies have been chiefly aimed at checking the theory, we here take another position: Assuming the current theory of quantum electrodynamics, one can combine the calculated and measured fine-structure intervals of helium in order to determine the value of the fine-structure constant α. At present, the best value of α comes from the work of Dayhoff, Triebwasser, and Lamb[5] on the fine structure $(2P_{1/2}-2P_{3/2})$ of hydrogen. That value is uncertain to about one part in 10^5, due essentially to the short lifetime of the $2p$ state, and it has not appeared feasible to improve on those measurements. However, the lifetime of the ^3P state of helium is about 2 orders of magnitude longer, and it is expected[6] that these fine structure intervals can be measured to an accuracy of one part in 10^6 or perhaps better.[7] It then becomes necessary to calculate theoretical values for the helium fine structure to $1/10^6$, and this is the task we here commence. It may be remarked that in the case of the hydrogen fine structure the theoretical formula[8] is very simple; this is not the case for helium. However, we believe that those circumstances which made the hydrogen problem simple, will have the effect of reducing the helium calculation from "impossibly difficult" to merely "difficult."

We should add that there is at present a specific need for a better value of α. Analysis[9] of the hyperfine structure of hydrogen reveals effects due to the electro-

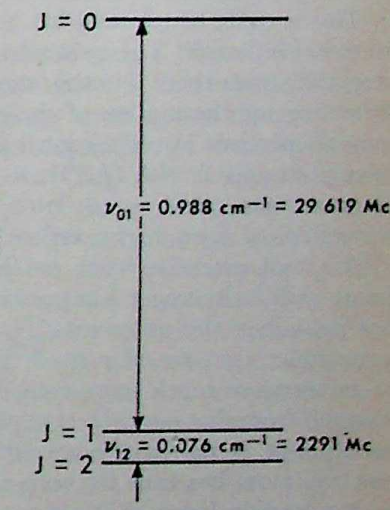


FIG. 1. Energy levels of the $1s2p\,^3P$ state of helium.

* This work was supported in part by the U. S. Atomic Energy Commission and the U. S. Air Force under Grant AF-AFOSR 130-63, and in part by the Advanced Research Projects Administration through the U. S. Office of Naval Research.

† A preliminary report of this work was presented at the 1963 Annual Meeting of the American Physical Society, Bull. Am. Phys. Soc. 8, 20 (1963).

[1] G. Breit, Phys. Rev. 36, 383 (1930). References to other early studies may be found in this reference.

[2] G. Araki, M. Ohta, and K. Mano, Phys. Rev. 116, 651 (1959), and several earlier papers referred to therein.

[3] J. Traub and H. M. Foley, Phys. Rev. 116, 914 (1959).

[4] C. L. Pekeris, B. Schiff, and H. Lifson, Phys. Rev. 126, 1057 (1962).

[5] E. S. Dayhoff, S. Triebwasser, and W. E. Lamb, Phys. Rev. 89, 106 (1953).

[6] V. W. Hughes (private communication).

[7] For the best measurements to date (accurate to about $0.5/10^6$), see J. Lifsitz and R. H. Sands, Bull. Am. Phys. Soc. 6, 424 (1961), and earlier references given therein.

[8] See, for example, H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press Inc., New York, 1957), p. 105.

[9] C. K. Iddings and P. M. Platzman, Phys. Rev. 115, 919 (1959); also see D. E. Zwanziger, Bull. Am. Phys. Soc. 6, 514 1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

magnetic structure of the proton, but a clear interpretation of those results is spoiled by the present uncertainty in the value of α.

II. PLAN OF THE CALCULATIONS

We will now outline the several parts of the task of calculating the fine structure intervals in helium to an accuracy of about one part in 10^6. The zeroth-order problem is the nonrelativistic two-electron Hamiltonian in a fixed Coulomb field:

$$H_0=\frac{p_1^2}{2m}+\frac{p_2^2}{2m}-\frac{Ze^2}{r_1}-\frac{Ze^2}{r_2}+\frac{e^2}{r_{12}}. \tag{1}$$

The next terms in a nonrelativistic expansion are the well-known[10] fine-structure terms which we shall denote as α^2H_2. [Our reference for energy is $e^2/a_0=\alpha^2mc^2$, and we shall not note factors of Z $(=2)$ as being pertinent to the classification of the expansion terms.] We thus expect an energy level formula something like

$$E_J-E_0=\alpha^2\langle H_2\rangle_J+\alpha^4\left[\langle H_4\rangle_J+\left\langle H_2\frac{1}{E_0-H_0}H_2\right\rangle_J\right]$$
$$+(\text{terms of order }\alpha^5\text{ and smaller which we shall ignore}), \tag{2}$$

where H_4 is some higher order operator which has not yet been worked out.

There are still many terms not represented in this formula (2). The most apparent is the correction for the anomalous magnetic moment of the electron, of order α^3. This is quite simple and has already been included by some authors.[2,4] The reduced mass must be put in properly, and there is also the operator $\mathbf{p}_1\cdot\mathbf{p}_2/M$ correcting for the motion of the nucleus; this is itself spin-independent but will contribute in a second-order calculation mixed with H_2. There are then very many diagrams one could write down describing radiative corrections of many higher orders.

The most extensive work on the fine structure and Lamb shift in hydrogen has produced terms up to, but not including, the order α^7mc^2,[11] which is just to the same order that we now need. But the two-electron atom seems so much more complicated than the one-electron atom that our project might appear too difficult to attempt. A thorough relativistic theory of the two-electron atom has thus far been analyzed[12] only as far as the leading Lamb-shift terms, $\alpha^5mc^2\times(\ln\alpha+\text{const})$, and our goal is well beyond this. What encourages us to proceed is the expectation that most of the difficult terms at these higher orders are spin-independent, and thus, while affecting the absolute energy level, do not contribute to the fine structure.

As an example, consider the effect on the fine structure of some deviation from pure Coulomb potentials occurring at some small distance R. The integrals which occur are of the general form of the familiar $\langle(1/r)\times(\partial V/\partial r)\rangle\sim\langle 1/r^3\rangle$, and are taken over radial wave functions for $l>0$ orbitals. Thus the relative correction would appear to be of the order

$$\int_0^R r^2dr r^{2l-3}\Big/\int_0^a r^2dr r^{2l-3}\approx(R/a)^{2l}\leqq(R/a)^2.$$

Thus the effect of finite nuclear size would seem to be completely negligible, and the effect of vacuum polarization would appear to be of order $\alpha(\alpha)^2=\alpha^3$ on the fine structure intervals, again negligible.

The entire project may be divided into three rather well-separated parts. First, is the task of constructing a sufficiently accurate eigenfunction of H_0 so that the leading term in Eq. (2) can be determined to about $1/10^6$; this is mostly a computer problem. Second, is the task of carrying out a proper relativistic analysis and determining the operator H_4 and whatever else may appear to the required order; this will be a matter needing both formal technique and enlightened short-cutting. The third job will be the evaluation of these smaller corrections to the fine structure, and this will require a combination of modest computer effort and much algebraic detail.

The rest of this paper is concerned with the first of these tasks.

III. THE SCHRÖDINGER WAVE FUNCTION: FIRST ATTEMPT

We wish to construct a good approximation to the lowest ^3P eigenfunction of the equation

$$H_0\psi=E_0\psi. \tag{3}$$

This will be done by setting up a sequence of trial functions ψ, using the variational principle

$$\delta\langle\psi|H_0-E_0|\psi\rangle=0 \tag{4}$$

and seeing how results of interest converge as we make the trial functions larger and larger.

We started with the functions

$$\psi=\left(\frac{1-P_{12}}{4\pi\sqrt{2}}\right)\sum_{l,m,n=0}^{l+m+n\leqq\omega}C_{lmn}\mathbf{r}_1r_1^mr_2^nr_{12}^l\times e^{-(\kappa\sigma/2)r_1}e^{-(\kappa/2)r_2}, \tag{5}$$

where P_{12} exchanges coordinated r_1 and r_2, and the P-state character is represented by the vector sign. The unit of length is $a_0=\hbar^2/me^2$; and the scale parameters κ and σ were taken, by extrapolation from the results of

[10] See Ref. 8, p. 181.

[11] A. J. Layzer, J. Math. Phys. 2, 308 (1961), has calculated terms of order $mc^2\alpha^7\ln^2\alpha$ and $mc^2\alpha^7\ln\alpha$. Only the latter contributes to the fine structure, and we hope it will not be difficult to include the term of this order in our problem.

[12] J. Sucher, Phys. Rev. **109**, 1010 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Result of variational calculations of $2\,^3P$ with the standard Hylleraas basis (5).

ω	Number of terms	Energy $-E(e^2/a_0)$	$C(\frac{1}{2}\alpha^2$ Ry)	$D(\frac{1}{2}\alpha^2$ Ry)
2	10	2.132678402	−0.07065132	−0.05404532
3	20	2.133085039	−0.06829756	−0.05379667
4	35	2.133140222	−0.06771029	−0.05394716
5	56	2.133157595	−0.06745614	−0.05400336
6	84	2.133162268	−0.06733560	−0.05402833
7	120	2.133163594	−0.06728399	−0.05403823
8	165	2.1331639812	−0.06726722	−0.05404284
9	220	2.1331641067	−0.06726320	−0.05404503
10	286	2.1331641530	−0.06726453	−0.05404627
Final result from Table II		2.1331641908	−0.06727529	−0.05404839

Traub and Foley,[3] to be 4.62 and 0.29,[13] respectively, and were not varied. The variation of the linear parameters C_{lmn} leaves us then with the numerical problem of finding the eigenvalue and eigenvector of a symmetric matrix of degree

$$N(\omega)=\tfrac{1}{6}(\omega+1)(\omega+2)(\omega+3). \tag{6}$$

Some of the details of the construction of the matrix elements and the numerical methods used will be found in the Appendix. We give in Table I the resulting eigenvalues for the series of trial functions of degree $\omega=2,\ 3,\ \cdots,\ 10$. This was done with 16-decimal arithmetic on an IBM-704 and consumed a total of some 20 h of machine time.

This work very closely parallels that of Pekeris, Schiff, and Lifson,[4] differing only in the specification of the exponential parameters, κ and σ. This makes a non-negligible difference: At 220 terms our eigenvalue is an order of magnitude better converged than theirs.[14]

Our chief interest anyway is not the energy eigenvalue, but the eigenvector, or more properly the average values of the two spin-dependent parts of H_2. These are the well-known[10] spin-orbit terms

$$H(\text{s.o.})=\left(\frac{e\hbar}{2mc}\right)^2\left(\frac{\boldsymbol{\sigma}_1+\boldsymbol{\sigma}_2}{2}\right)\cdot\left[Z\left(\frac{\mathbf{r}_1\times\mathbf{p}_1}{r_1^3}+\frac{\mathbf{r}_2\times\mathbf{p}_2}{r_2^3}\right)-\frac{3(\mathbf{r}_1-\mathbf{r}_2)\times(\mathbf{p}_1-\mathbf{p}_2)}{r_{12}^3}\right], \tag{7a}$$

and the spin-spin term

$$H(\text{s.s.})=\left(\frac{e\hbar}{2mc}\right)^2\frac{1}{r_{12}^3}\left(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2-\frac{3\boldsymbol{\sigma}_1\cdot\mathbf{r}_{12}\boldsymbol{\sigma}_2\cdot\mathbf{r}_{12}}{r_{12}^2}\right). \tag{7b}$$

[We have dropped here a term in H(s.o.) which has only off-diagonal matrix elements.] The first-order fine-structure levels are customarily given in terms of two constants C (from s.o.) and D (from s.s.):

$$\begin{aligned}\langle H_2\rangle_{J=0}&=\text{const}-2C-(10/3)D,\\ \langle H_2\rangle_{J=1}&=\text{const}-C+(5/3)D,\\ \langle H_2\rangle_{J=2}&=\text{const}+C-(1/3)D.\end{aligned} \tag{8}$$

[These values will subsequently be given in units of $(e\hbar/2mc)^2a_0^{-3}=(\alpha^2/4)(e^2/a_0)=\frac{1}{2}\alpha^2$ Ry]. These results are also shown in Table I. The accuracy of our C and D values here is hardly better than that of Pekeris *et al.*[4] (about $1/10^4$) and far short of the desired $1/10^6$. At the time when these results were obtained we attempted to extrapolate the apparent rate of convergence of the output numbers, and thus estimated that something approaching 2000 terms of the series (5) would be needed to obtain the required accuracy in C and D. It seemed that, at best, this would be an extremely expensive venture; and we thus decided that this attack had failed.

Two possible paths were then considered. We might start with the wave functions already constructed, and, by use of auxiliary variational calculations of modest size, attempt to increase the accuracy of the C and D integrals to something approaching that which is obtained directly for the energy eigenvalue.[15] However, it appeared that, due to the rather singular nature of the operators H_2, the auxiliary functions needed for this method would be of a complicated analytical nature; and so we did not make a serious attempt in this direction.

Alternatively, we could seek a better set of basis functions for the calculation of ψ. The most likely cause for the slow convergence noted above appeared to be the weak logarithmic singularity in the two-electron wave function studied by Fock.[16,17] It appeared, however, that putting into ψ the explicit logarithm term given by Fock would be a very messy job. Furthermore, it was felt that the reward would be rather slight if only the first term were accounted for; and the higher terms of Fock's expansion appear to be complicated beyond our ability to manage. We then sought, in a much more *ad hoc* manner to introduce just any covenient terms which would add flexibility to ψ especially in the region ($r_1\to 0$ and $r_2\to 0$). We thus chose to try the addition of the simple factor $(r_1+r_2)^{1/2}$ to the series (5).[18] The efficacy of this extension of the basis functions was first tested on the relatively simple calculation of the

[13] The precise values of κ and σ used were about 10^{-8} smaller than the figures given here, due to the decimal-to-binary conversion operation of the machine.

[14] For a discussion of relative convergence rates, see the article by C. Schwartz in *Methods in Computational Physics* (Academic Press Inc., New York, 1963), Vol. 2, p. 241.

[15] The method referred to is that described by C. Schwartz, Ann. Phys. (N.Y.) **6**, 170 (1959).

[16] V. A. Fock, Izv. Akad. Nauk SSSR Ser. Fiz. **18**, 161 (1954).

[17] See also the discussion and analysis in Ref. 14, p. 256.

[18] Half powers were earlier introduced into the Hylleraas expansion by H. M. Schwartz, Phys. Rev. **120**, 483 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Results of variational calculations of $2\,^3P$ with the extended basis (9).

Number ω	of terms	$-E(e^2/a_0)$	In units of $\frac{1}{2}\alpha^2$ Ry Cz	C_e	D
1	7	2.132580318470	0.127974449	−0.198448086	−0.053285232
2	19	2.133100285189	0.136827714	−0.203765520	−0.053888850
3	39	2.133155369966	0.138223897	−0.205167872	−0.053987664
4	69	2.133162942721	0.138649978	−0.205707289	−0.054026808
5	111	2.133164012406	0.138640520	−0.205858266	−0.054041649
6	167	2.133164164417	0.138638567	−0.205898542	−0.054046487
7	239	2.133164186301	0.138635834	−0.205908639	−0.054047869
8	329	2.133164189955	0.138636386	−0.205911292	−0.054048248
9	439	2.133164190626	0.138636755	−0.205911984	−0.054048351
Extrap-olation		2.13316419080 ±5	$C=Cz+C_e\rightarrow$	−0.067275287 ±19	−0.054048390 ±13

helium ground state (1^1S), and those very successful results have already been presented.[19]

The calculations reported in this section are not a total loss, since we can expect to use these medium-accuracy wave functions for the later calculation of the small correction terms schematically represented in Eq. (2).

IV. THE WAVE FUNCTION: SECOND ATTEMPT

The expanded basis now looks like the expression (5) with the replacement

$$C_{lmn}\rightarrow C_{lmn}+D_{lmn}(r_1+r_2)^{1/2}; \tag{9}$$

and the single term D_{000} is omitted. The scale parameter σ was left fixed as before at 0.29,[13] but after a short search we changed the value of κ to 4.0 (exactly). The matrix elements of H_0 and H_2 are more complicated with the half-power terms, but once we had learned to evaluate the integrals C and D even with the old basis, the new techniques were not very difficult (see Appendix). The results converged much more rapidly, than before, but it was found, as might be expected that as one approached more closely to the exact fitting of the function numerical accuracy became more and more critical at an alarming rate. We found it necessary to construct special programs to do all our arithmetic to an accuracy of 30, and then 52 decimals; the resulting cost in machine time was very great. The total labor took some 30 h on an IBM-7090.

In Table II are presented our results with the half-power functions up to the ninth order. We show here separately the contributions from the two parts of (7a):

$$C=C_Z+C_e. \tag{10}$$

V. RESULTS

From the numbers in Table II we now must give our best values for the fine-structure intervals along with an estimate of their probable uncertainty. This we do by attempting to extrapolate the results shown to $\omega\rightarrow\infty$; the smoothness of the sequence of computed numbers will be used to indicate the reliability we may place on this extrapolation.

[19] C. Schwartz, Phys. Rev. 128, 1146 (1962).

TABLE III. Extrapolation of the calculated interval ν_{12} (in units of $\frac{1}{2}\alpha^2$ Ry).

Number of terms	Calculated $2D-2C$	Successive differences	Ratios of successive differences
7	0.0343768105		
		−0.008279	
19	0.0260979112		0.022
		−0.00018529	
39	0.0259126200		−0.800
		0.0001483866	
69	0.0260610066		1.962
		0.0002911883	
111	0.0263521949		0.257
		0.0000747812	
167	0.0264269761		0.306
		0.0000228957	
239	0.0264498718		0.151
		0.0000034446	
329	0.0264533164		0.127
		0.0000004379	
439	0.0264537543		
	Projected increment = 0.000000064 ±21		
	Extrapolation = 0.026453818 ±21		

Table III shows the details of such an extrapolation for the small interval

$$\nu_{12}=\tfrac{1}{2}\alpha^2\ \text{Ry}\cdot(2D-2C). \tag{11}$$

(It does make a helpful difference on the final error estimates that we extrapolate the combined integrals, rather than combine the extrapolated integrals.) The table shows values, differences, and ratios of successive differences; attempts to extrapolate must be based on the behavior of these ratios, especially at the bottom of the sequence. It is clear from Table III that these ratios cannot be very well described in any simple analytical terms; but it is to our advantage that they are very small.[20] After some playing with these numbers, attempting various analytical and graphical fits, it was decided that a reasonable procedure would be simply to extrapolate from the last step as if we had a geometric series, and assign as our uncertainty $\frac{1}{3}$ of the added increment. That is, if Δ' is the last difference recorded and R' is the last ratio, our final answer is gotten by adding to the last computed value the quantity

$$\Delta''=\Delta'\left(\frac{R'}{1-R'}\right)(1\pm\tfrac{1}{3}). \tag{12}$$

[20] By way of contrast one may note that the corresponding ratios in the work of Pekeris *et al.* (Ref. 4) are much smoother, but also quite a bit larger. One may also notice that while their value for ν_{01} calculated with 220 terms is 0.011% larger than our final result, their extrapolated value is 0.009% too small. Their extrapolation here was thus in error by about 45%. (For ν_{12} their extrapolation erred by only about 16%) We point these numbers out intending not to chasten others, but to emphasize the need for a critical attitude toward any attempts at numerical extrapolation, and in particular our own.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We thus arrive at the following results for the intervals of the 2^3P state of helium, bare of any higher order corrections whatsoever.

$$\nu_{01}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.33751721(7)]$$
$$R'=0.210 \tag{13a}$$

$$\nu_{12}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.02645382(2)]$$
$$R'=0.127 \tag{13b}$$

$$\nu_{02}=\tfrac{1}{2}\alpha^2\ \mathrm{Ry}[0.36397101(9)]$$
$$R'=0.172. \tag{13c}$$

The number in the inner parenthesis gives the uncertainty in the last figure quoted, arrived at according to (12). If our extrapolations are accepted, then we have achieved the desired goal of $1/10^6$ accuracy. If it becomes necessary, we could, without too much expense, carry the numerical work one order further and thus increase the certainty of these numbers by about a factor of 5; but at present we feel this part of the program can rest, and wait for the experimental work and the analysis of higher-order corrections to catch up.

VI. CONCLUSION

The first major goal of the plan outlined in Sec. II has now apparently been achieved, and this encourages us to proceed with the other parts of the over-all program for the redetermination of α. There are two rather apparent cautions which we wish to make. First, regarding the extrapolation and error estimation, we have tried to be both clever and objective in our analysis, but each reader will have to find his own measure of skepticism regarding the precise accuracy of our results. Secondly, while we have endeavored to check and recheck all our algebra and computer programs, the considerable complexity of the work reported here must leave open the possibility of some undiscovered error of importance. To this point we report the following experience.

Six months after we had completed the last of our computations, there was discovered an error in our computer program for multiple precision division. It turned out that the last ten decimals were sometimes treated incorrectly; but some conditions under which this error would not arise could also be precisely identified. A detailed scrutiny of all our programs then revealed the astonishing fact that this mistake did not have any effect at all on any of the matrix element computations. The error could work only in the process of solving for the eigenvectors of the matrices; but our method was, at the end, an iterative process, so that we suffered only a slower convergence (paid more dollars) as a result of the mistake, and all our final answers were unaffected.

We hope that someone, working independently of us, will repeat and check all this work; this is the best insurance we can imagine for the reliability of these important results.

One can ask for a comparison of the present best experimental and theoretical values for the fine-structure intervals. Adding to our results (13) the simple correction for the anomalous magnetic moment of the electron[2] (the only correction of relative order α) and using the value of the Rydberg for helium,

$$\tfrac{1}{2}\alpha^2\ \mathrm{Ry}=2.921374\ \mathrm{cm}^{-1}, \tag{14}$$

we get

$$\nu_{01}=0.987837\ \mathrm{cm}^{-1},$$
$$\nu_{12}=0.0765302\ \mathrm{cm}^{-1}, \tag{15}$$
$$\nu_{02}=1.064367\ \mathrm{cm}^{-1}.$$

These may be compared with the experimental values

$$\nu_{01}=0.98791(4)\ \mathrm{cm}^{-1},$$
$$\nu_{12}=0.076423(3)\ \mathrm{cm}^{-1},$$
$$\nu_{02}=1.06434(3)\ \mathrm{cm}^{-1}.$$

[*Note added in proof.* The numbers quoted here have been corrected, relative to those published (Ref. 7), according to private communication from Dr. R. H. Sands. See also F. M. J. Pichanick, R. D. Swift, and V. W. Hughes, Bull. Am. Phys. Soc. 9, 90 (1964).] The differences are about 0.0001 cm^{-1}, which is roughly what one would expect due to the neglected terms of relative order α^2. Araki[2] has considered one such higher-order term, the mixing with the nearby 2^1P state; Perkeris *et al.*[4] give for this effect a resulting downward shift of 0.000158 cm^{-1} to the $J=1$ level. The addition of this correction appears to do little more than change the signs of the discrepancies for the 0–1 and 1–2 intervals; and this indicated that all the higher order terms of Eq. (2) will have to be treated together.

ACKNOWLEDGMENTS

The author was assisted in some of the early programming by A. Miller. He also wishes to thank many staff members of the University of California Computer Center for the advice and service they have generously given him.

APPENDIX: NUMERICAL METHODS

The many matrix elements of the operators H_0 and H_2 in the basis used were evaluated in terms of pretabulated integrals of the following general type:

$$A_\Lambda(L,M,N)=\left(\frac{1+\sigma}{2\sigma}\right)^{L+M+N+(\Lambda/2)}\int\frac{dv_1}{4\pi}\int\frac{dv_2}{4\pi}e^{-r_1}e^{-r_2/\sigma}$$
$$\times r_1^{M-2}r_2^{N-2}r_{12}^{L-2}f_\Lambda(r_1,r_2), \tag{A1}$$

$B_\Lambda(L,M,N)$ = same as (A1) but with a factor

$$\cos\theta_{12}\ \text{under the integral.} \tag{A2}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

L, M, N are integers greater than zero; $\sigma(0<\sigma<1)$ is a general parameter which is put $=0.29$ for the direct integrals and $=1$ for the exchange integrals; and f_Λ is a homogeneous function of degree $\frac{1}{2}\Lambda$ in the two lengths r_1 and r_2.

By use of the following identity referring to integration over the angle θ_{12} (signified by the angular brackets)

$$\langle \sin^2\theta_{12} r_{12}{}^L\rangle = \frac{-2}{(L+2)}\left\langle \cos\theta_{12}\frac{r_{12}{}^{L+2}}{r_1 r_2}\right\rangle, \qquad \text{(A3)}$$

one can easily establish the following recursion formulas:

$$\begin{aligned} A(L,M,N) &= A(L-2,M+2,N)+A(L-2,M,N+2) \\ &\quad -2B(L-2,M+1,N+1), \\ B(L,M,N) &= [(L-2)/(L+2)][B(L-2,M+2,N) \\ &\quad +B(L-2,M,N+2) \\ &\quad -2A(L-2,M+1,N+1)]. \end{aligned} \qquad \text{(A4)}$$

Our general method for building tables of integrals is by the use of such recursion formulas; but one must be careful that the particular recursive scheme chosen does not build by taking differences of nearly equal numbers (most obvious recursion schemes do suffer from this serious drawback.) For the above integrals—for large values of L, M, N in particular—the main contribution comes from large values of r_{12}; thus $\cos\theta_{12}$ is mostly negative, the integrals A and B have opposite signs, and the recursion schemes (A4) are safe.

In order to start the recursion sequence we need to evaluate A and B for $L=1$ and 2. The angular integrals are simple and we can express the results in terms of two families of two-dimensional integrals. (The subscript Λ is understood in what follows.)

$$\begin{aligned} &B(2,M,N)=0, \\ &A(2,M,N)=F(M+1,N+1)+G(N+1,M+1), \\ &B(1,M,N)=\tfrac{1}{3}F(M+2,N-1)+\tfrac{1}{3}G(N+2,M-1), \\ &A(1,M,N)=F(M+1,N)+G(N+1,M), \end{aligned} \qquad \text{(A5)}$$

where

$$\begin{aligned} F(M,N) &= [(1+\sigma)/2\sigma]^{M+N+(\Lambda/2)}\int_0^\infty dr e^{-r} r^{M-1} \\ &\quad \times\int_r^\infty ds e^{-s/\sigma} s^{N-1} f_\Lambda(r,s) \\ &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\Gamma\left(M+N+\frac{\Lambda}{2}\right) \\ &\quad \times\int_0^1 dx \frac{x^{M-1} f_\Lambda(x,1)}{[(1/\sigma)+x]^{M+N+(\Lambda/2)}}, \end{aligned} \qquad \text{(A6)}$$

and

$$\begin{aligned} G(M,N) &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\int_0^\infty dr e^{-r/\sigma} r^{M-1} \\ &\quad \times\int_r^\infty ds e^{-s} s^{N-1} f_\Lambda(s,r) \\ &= \left(\frac{1+\sigma}{2\sigma}\right)^{M+N+(\Lambda/2)}\Gamma\left(M+N+\frac{\Lambda}{2}\right) \\ &\quad \times\int_0^1 dx \frac{x^{M-1} f_\Lambda(1,x)}{[1+(x/\sigma)]^{M+N+(\Lambda/2)}}. \end{aligned}$$

We shall again tabulate these two dimensional arrays, F and G, by recursion; but all the obvious formulas we have looked at (proceeding from small to large values of M and N) are badly behaved regarding loss of accuracy from subtractions. We thus use the backward iteration formulas

$$\begin{aligned} F(M,N) &= \frac{2/(1+\sigma)}{(M+N+\frac{1}{2}\Lambda)}[\sigma F(M+1,N)+F(M,N+1)] \\ G(M,N) &= \frac{2/(1+\sigma)}{(M+N+\frac{1}{2}\Lambda)}[G(M+1,N)+\sigma G(M,N+1)]. \end{aligned} \qquad \text{(A7)}$$

These recursions start from the values of F and G along the line $M+N=$ constant (about 30 in our work). We have found the following devious, but numerically safe, procedure for these evaluations. After Eq. (A6) make the change of variables

$$x=\frac{1-u}{1+u}, \qquad \text{(A8)}$$

implying

$$\begin{aligned} f_\Lambda(x,1) &= (1+u)^{-(\Lambda/2)} f_\Lambda(1-u,1+u), \\ f_\Lambda(1,x) &= (1+u)^{-(\Lambda/2)} f_\Lambda(1+u,1-u); \end{aligned} \qquad \text{(A8)}$$

then expand the resulting denominator of F in an infinite power series in

$$[\tfrac{1}{2}(1-\sigma)(1-u)],$$

and the denominator of G in terms of $\{[(1-\sigma)/(1+\sigma)]\times u\}$. Also make the binomial expansion of the resulting numerator factor $(1+u)^{N-1}$ according to the separation

$$(1+u)=2u+(1-u).$$

The resulting expressions are

$$\begin{Bmatrix} F(M,N) \\ G(M,N) \end{Bmatrix} = \sum_{J=1}^{N} \frac{(N-1)!2^{J-1}(M+N-J-1)!}{(N-J)!2^{M+N+(\Lambda/2)-1}} \times \begin{Bmatrix} \tilde{F}_J(M+N) \\ \tilde{G}_J(M+N) \end{Bmatrix}, \qquad \text{(A9)}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$\tilde{F}_J(K)=[\tfrac{1}{2}(1+\sigma)]^{K+(\Lambda/2)}\sum_{k=0}^{\infty}\left(\frac{1-\sigma}{2}\right)^k \times\frac{(K+\frac{1}{2}\Lambda-1+k)!}{k!(J-1)!(K-1-J)!}\tilde{F}_J'(k,K), \tag{A10}$$

$$\tilde{G}_J(K)=\sum_{k=0}^{\infty}\left(\frac{1-\sigma}{1+\sigma}\right)^k\frac{(K+\frac{1}{2}\Lambda-1+k)!}{k!(J-1)!(K-1-J)!}\tilde{G}_J'(k,K),$$

and

$$\tilde{F}_J'(k,K)=\int_0^1 du f_\Lambda(1-u,1+u)u^{J-1}(1-u)^{K-1-J+k}, \tag{A11}$$

$$\tilde{G}_J'(k,K)=\int_0^1 du f_\Lambda(1+u,1-u)u^{J-1+k}(1-u)^{K-1-J}.$$

For most of the integrals of interest the function $f_\Lambda(r_1,r_2)$ is $(r_1+r_2)^{\Lambda/2}$ ($\Lambda=0$, ±1, or 2); and so in the final expressions (A11) the function f_Λ is merely a constant and these integrals are trivial. However in treating H_2 we also run into the spectral case of Eq. (A3) for $L=-2$, and this leaves us with

$$f_\Lambda(r_1,r_2)=\ln\left|\frac{r_1-r_2}{r_1+r_2}\right|\cdot(r_1+r_2)^{\Lambda/2} \tag{A12}$$

in Eq. (A6); but this reduces to $2^{\Lambda/2}\ln u$ in Eq. (A11), and these final integrals are again easy.

The infinite sums in Eq. (A10) are rapidly converging for our values of σ; and the entire procedure worked very well.

Matrices of dimension 440 in sextuple precision require over half a million words of machine storage space (an order of magnitude more than what is available in the core) and so we made extensive use of magnetic tapes. The most time-consuming operation was the matrix inversion carried out at the start of the eigenvector calculation. For the largest, dimension 438, this took 10 h of 7090 time—equivalent to about one billion simple 8-figure multiplications on this machine.

We will now describe the method used for finding the eigenvector. Given two symmetric matrices H and N ($N>0$) of order n, we separate the problem

$$\sum_{j=1}^{n}(H_{ij}-\lambda N_{ij})x_j=0,\quad i=1,2,\cdots,n \tag{A13}$$

as follows: Set $x_1=1$, and solve for x_i from

$$\sum_{j=2}^{n}(H_{ij}-\lambda N_{ij})x_j=-(H_{i1}-\lambda N_{i1}),\ i=2,3,\cdots,n \tag{A14}$$

once λ is known; but we can get a stationary estimation for λ, once an approximate vector x is known, from the Rayleigh quotient,

$$\lambda=(x,Hx)/(x,Nx). \tag{A15}$$

In this work we always have a very good guess for λ to start with, and the alternating iteration (A14), (A15) is very rapidly convergent. Most of the time goes into solving (A14); the process of solving n simultaneous linear equations

$$Ax=b \tag{A16}$$

by the direct method of elimination takes $\frac{1}{6}n^3$ operations (for a symmetric matrix A). On the other hand, iterative attacks on the solution of (A16) require only of the order of n^2 operations per cycle, but many iteration schemes do not converge at all well. The general criterion is that one have some "good" approximation to the inverse of the matrix A; then

$$x^{(l+1)}=x^{(l)}+\tilde{A}\cdot(b-Ax^{(l)}), \tag{A17}$$

where $\tilde{A}$ approximates A^{-1}, can be expected to converge "rapidly." For our problem the matrix to be inverted (A14) changes only very slightly, as our value of λ is improved, so that it is necessary to calculate *only once* an inverse matrix (by the direct method[21]), and then we use this as the kernel of the iterations (A17), until both eigenvalue and eigenvector are converged.

This process works well, but as one goes to larger and larger matrices, loss of numerical accuracy becomes a serious problem at a rate much beyond any that could be attributed to statistical phenomena. Obviously, the full matrix $(H-\lambda N)$ in Eq. (A13) is singular—for the correct λ—but one would think that the remnant matrix in Eq. (A14) is far from singular since the most important element of the basis has been removed. In actuality, however, the removal of a single trial function does not prevent the convergence of the variational sequence, it may merely slow it. Thus, as we go to larger and larger bases, the matrix in Eq. (A14) does come closer and closer to being singular, and at a painfully fast rate. We handled this problem by the brute force method of using higher precision arithmetic, but the cost was quite great. (It may be that the problem would not have been quite as bad as it did appear to us, had we caught earlier the error mentioned in Sec. VI.)

[21] After completion of this work the following, possibly more efficient, method occurred to us. The new terms one is adding to the basis at each step in the sequence of variational calculations must be numerically of decreasing importance (if the entire process is converging at all well). Then it may be sufficient to get $(\tilde{A})$ by carrying out the direct calculation on only a small submatrix representing the dominant terms, and simply using the diagonal elements for the higher parts of the matrix.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Low-Energy Expansion of the Scattering Amplitude for Long-Range Quadrupole Potentials*

THOMAS F. O'MALLEY†
Joint Institute for Laboratory Astrophysics, Boulder, Colorado

(Received 20 January 1964)

For static potentials which are proportional asymptotically to $qP_2(\cos\theta)/r^3$, the low-energy expansion of the scattering amplitude is found through terms of $O(k)$, using a modification of the method developed by Levy and Keller for central potentials. The resulting expansion to lowest order in k is found to be $f(\theta,\phi)\sim -A + (q/3)P_2(\cos\theta_K)+O(k)$, where A is the scattering length and θ_K is the coordinate of the momentum transfer vector. Applications are attempted first to electron-atom elastic scattering where results are somewhat more complicated than for the potentials above, secondly to transitions between magnetic quantum states of atoms caused by slow electrons.

1. INTRODUCTION

EFFECTIVE range expansions developed originally for short-ranged potentials[1] have been extended in recent years first to spherical potentials with r^{-4} tails[2] and finally to potentials with arbitrary r^{-n} radial dependence asymptotically.[3] This has enabled the usefulness of the low-energy expansion to be extended to electron-atom collisions,[4] and to any scattering problem where long-range Van der Waals forces are operative. However there are physical problems, exemplified by the scattering of electrons from atomic oxygen or other nonsymmetric atoms, for which the asymptotic form of the interaction is not the r^{-4} polarization potential but rather the quadrupole interaction proportional to $r^{-3}P_2(\cos\theta)$; in particular:

$$(2m/\hbar^2)V(\mathbf{r})\sim qP_2(\cos\theta)/r^3-r^{-4}[p+p'P_2(\cos\theta)] +O(r^{-5}). \quad (1.1)$$

Since such a potential is not spherically symmetric, the usual phase-shift analysis is not applicable for one has a coupling of different angular momentum states. Hence the expansions of Levy and Keller[3] for the phase shifts are not applicable. But while phase shifts for this problem are not well defined, the scattering amplitude, $f(\theta,\phi)$, is. In this paper we will look for an energy expansion for $f(\theta,\phi)$,[5] adapting to this purpose the approach which was used by Levy and Keller[3] (LK) to expand $\tan\eta$ for symmetric potentials.

In Sec. 2, we shall briefly review the LK method, modifying it to a more heuristic form which we find easier to apply to the present purpose. In Sec. 3, the method is then applied to expanding the amplitude, $f(\theta,\phi)$, for potentials with the asymptotic form (1.1). The dipole potential is also briefly considered. Section 4 is devoted to considering the relation of the results of the previous section to the electron-atom scattering problem. In particular, the derivation of the form (1.1) of the potential is indicated and the relation of q to the atom's quadrupole moment is discussed. It is seen that that electron-atom problem is somewhat more complicated, due to the possibility of atomic transitions from one magnetic quantum state to another. The amplitude for these transitions is obtained to lowest order in the energy.

2. MODIFIED LEVY–KELLER METHOD

To introduce the method which will be applied in Sec. 3, consider the $L=0$ scattering of a particle by a long-range central potential, $U(r)$, proportional to r^{-n} asymptotically. Levy and Keller's starting point in attacking this problem was the first-order nonlinear equation[6] for the phase shift

$$d/dr\,\tan\eta(r)=-k^{-1}U(r) \times[\sin(kr)+\tan\eta(r)\cos(kr)]^2 \quad (2.1)$$

with boundary condition $\eta(0)=0$. $\eta(\infty)$ is the desired phase shift. They then represented $\tan\eta$ as a series in odd powers of k

$$\tan\eta(r)=k\sum_{j=0}^{\infty}k^{2j}i_j(r). \quad (2.2)$$

Equation (2.2) is substituted into (2.1) and equating like powers of k gives the equations for each of the

---

* A preliminary report of this work was given at the Winter Meeting in the West, Bull. Am. Phys. Soc. 8, 608 (1963).

† Present address: Defense Research Corporation, Santa Barbara, California.

[1] J. M. Blatt and J. D. Jackson, Phys. Rev. 76, 18 (1949); H. A. Bethe, *ibid.* 76, 38 (1949); G. F. Chew and M. L. Goldberger, *ibid.* 75, 1637 (1949).

[2] T. F. O'Malley, L. Spruch, and L. Rosenberg, J. Math. Phys. 2, 491 (1961).

[3] B. R. Levy and J. B. Keller, J. Math. Phys. 4, 54 (1963).

[4] T. F. O'Malley, L. Rosenberg, and L. Spruch, Phys. Rev. 125, 1300 (1962); T. F. O'Malley, *ibid.* 130, 1020 (1963).

[5] An alternate approach might be to look for an expansion of the elements of the reactance matrix connecting states of different angular momentum. If one is interested in the scattering amplitude only, this approach is more complicated. But as other information in addition to the scattering amplitude can be found this way, work along this line is being done.

[6] For a discussion of this equation, see Larry Spruch, in *Lectures in Theoretical Physics*, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience Publishers, Inc., New York, 1962), Vol. 4, p. 205 ff.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$l_j(r)$. $l_0(\infty)$ is the negative of the scattering length, A. Its equation is found to be

$$dA/dr = U(r)[r - A(r)]^2, \tag{2.3}$$

where we have changed the symbol l_0 to $-A$. The equations for l_1, etc., are similarly found.[3] If the potential were of short range, the problem would be finished, with the $l_j(\infty)$ furnishing the desired expansion of $\tan\eta$. However for long-range potentials, which are proportional asymptotically to r^{-n}, it is found that at most the first few of the l_j tend to a finite limit as $r \to \infty$. For example, if $n=4$, only $l_0(\infty)$ exists. [It should be pointed out for later reference that all the quantities $l_j(r)$ are still well defined for any finite r, and the expansion, Eq. (2.2) is still meaningful. One is merely prevented from going to the limit.]

Levy and Keller's procedure at this point was to consider the difference,

$$\tan\eta(r) - k\sum_{j=0}^{m} k^{2j} l_j(r) = k\beta(r,k),$$

where l_m is the last one whose limit is finite. Upon substituting this into (2.1), an equation for the function β was derived. This equation was then solved by an iteration or perturbation approach. It was found that, to lowest order in k, β was determined by the asymptotic region $r \to \infty$. In this region, the usefulness of the iteration method arises from the fact that successive iterations produce successively higher powers of k, so that a series in powers of k tends to fall out immediately.

For the purposes of this paper, a somewhat heuristic version of this method will be employed. We shall use the perturbation approach from the very beginning, and work directly with $\tan\eta$, rather than introduce the difference function, $\beta(r)$. The result of this is a greater directness and brevity, and a more intuitive approach. This will be done at the expense of introducing certain anomalous terms into the results, which however we shall see to have a clear interpretation.

To begin, let us rewrite Eq. (2.1) in integral form

$$\tan\eta(r) = \tan\eta(R) - k^{-1} \times \int_R^r U(r)[\sin(kr) + \tan\eta(r)\cos(kr)]^2. \tag{2.4}$$

R will be taken to be any very large but fixed distance (independent of k) and we vary k so that $kR \to 0$. The region $r < R$ is described by Eq. (2.2), as already pointed out. Now since R may be taken to be arbitrarily large, the integral may be considered to be a very small quantity for any allowable potential. (It is shown below that this imposes the mild restriction that r^2U approach zero.) Thus the aptness of a perturbation approach to Eq. (2.4) for $r > R$ is clear.

Since R is arbitrarily large, we substitute for $U(r)$ its asymptotic expansion,

$$U(r) \sim \sum_n U_n r^{-n}.$$

We further introduce the new variable $kr = x$. Equation (2.4) then becomes

$$\tan\eta(r) = \tan\eta(R) - \sum_n k^{n-2} U_n \times \int_{kR}^{kr} dx\, x^{-n}[\sin(x) + \tan\eta\cos(x)]^2. \tag{2.5}$$

If we temporarily ignore the k dependence of the lower limit of integration, the way in which the various terms in the asymptotic expansion of U contribute to the energy expansion of $\tan\eta$ follows by inspection from Eq. (2.5). In applying perturbation theory, one first takes $\tan(R)$ as the zeroth-order quantity and substitutes this for $\tan\eta$ on the right-hand side of Eq. (2.5). From Eq. (2.2), we see that this quantity is $O(k)$. In first-order perturbation theory then, each r^{-n} term in U contributes a k^{n-2} term to $\tan\eta$ plus terms of higher order. Now when these first-order terms are substituted back into the integral on the right-hand side, they produce in second order an additional factor k^{n-2}; and so provided the lowest n, n_0, is greater than 2, the leading term in second-order perturbation theory is $O[k^{2(n_0-2)}]$, and so on, just as found in LK. It is in this property that the power of the method lies. However, because of the k dependence of the lower limit of integration, there will be additional lower order terms in the total result.

We now proceed to study these terms and illustrate the method at the same time by solving for $\tan\eta(\infty)$ in first order, for a potential with $n_0=4$, the case which LK treated in detail. We assume for convenience that there are no other terms in the asymptotic expansion, and look for terms of order less than k^3. It will be seen that all the terms found by LK are reproduced very simply. In the zeroth order, we set for $r \geq R$

$$\tan\eta^{(0)}(r) = \tan\eta(R) = -A(R)k + O(k^3), \tag{2.6}$$

where the second equality follows from Eq. (2.2) with the notation $A = -l_0$. (This equation, as mentioned above, is valid up to any R which is fixed and finite.) Substituting this in the approriate places on the right-hand side of Eq. (2.5) gives for the first-order result

$$\begin{aligned}\tan\eta^{(1)}(\infty) = &-A(R)k - U_4k^2 \\ &\times \int_{kR}^{\infty} [\sin x - A(R)k\cos x]^2 \\ &\times x^{-4}dx + O(k^3) \\ = &-\left[A(R) + \frac{U_4}{R} - \frac{U_4A(R)}{R^2} \right. \\ &\left. + \frac{U_4A^2(R)}{3R^3}\right]k + \tfrac{1}{3}\pi U_4k^2 \\ &+ \tfrac{4}{3}U_4A(R)k^3\ln k + O(k^3). \end{aligned} \tag{2.7}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The integrals, which are elementary, are simplified if one notices that with the exception of the k^2 term all other terms come from the singularities at the lower limit of integration. [In all this analysis, since R is fixed, the quantity kR is considered an infinitesimal.] If one compares the result, Eq. (2.7), with the corresponding result Eq. (35) of LK[3] which is likewise first order, one sees that they are essentially the same. The only difference is that their η_0 (the negative of the scattering length) is replaced by the bracket term in the present result. A moment's reflection shows in what sense they are the same, namely the bracket represents the beginning of a perturbation expansion of A, starting with $A^{(0)}=A(R)$ in Eq. (2.3). This was to be expected since we are using a perturbation method throughout. It can be verified explicitly if one looks for the solution of Eq. (2.3) in first-order perturbation theory with the present potential. This is given by

$$A^{(1)}(\infty)=A(R)+U_4\int_R^\infty [r-A(R)]^2 r^{-4}dr$$

which when evaluated is exactly equal to the bracketed quantity in Eq. (2.7). These perturbation corrections to $A(R)$ are easily recognized in a calculation like Eq. (2.7) first by their k dependence and secondly by their explicit dependence on R. What they amount to is a renormalization of our zeroth-order scattering length. [If we had kept higher powers of k, there would have been renormalizations of the effective range as well.] The other terms in Eq. (2.7), the k^2 and $k^3 \ln k$ terms, are quite independent of the choice of R. Our practice in applying the method hereafter will be to replace, in all final results, both terms like the bracket term in Eq. (2.7) and quantities like $A(R)$ in the log term by the fully renormalized quantity $A=A(\infty)$, to which they tend as either $R\to\infty$ or the order of perturbation theory is increased.

Since the method just described is to be used in the following section to find the expansion of the scattering amplitude, a brief review of the procedure may be in order. One first finds the exact first-order nonlinear equation for the quantity f of interest corresponding to Eq. (2.1) above and writes it in integral form as in Eq. (2.4). One then considers the region r less than some large but fixed R, and makes the expansion corresponding to Eq. (2.2) for $f(R)$. Since this is the same whether the potential is short or long ranged, the form of the expansion will usually be known already. However, it is very simply derived by substituting an assumed expansion such as Eq. (2.2) into the exact equation and equating powers of k. Having done this, one is ready to study the contribution from $r>R$ by the perturbation or iteration method. R is assumed very large, the asumptotic expansion of $U(r)$ is substituted into the equation, and the new variable $x=kr$ is introduced. With $f(R)$ taken as the zero-order quantity, the iteration is begun. One then notices that the leading power of k introduced by each succeeding order of perturbation theory is of higher order in k than the last (apart from the renormalizations already discussed), and uses this to decide at which order to stop. Finally, one replaces partially renormalized quantities like the bracket and $A(R)$ in Eq. (2.7) by the fully renormalized quantity, such as A.

3. EXPANSION OF THE SCATTERING AMPLITUDE, $f(\theta,\phi)$

In this section we are interested in the solution of the Schrödinger equation

$$[\nabla^2+k^2-U(\mathbf{r})]\psi(\mathbf{r})=0$$

with the asymptotic form

$$\psi\sim\exp(i\mathbf{k}_i\cdot\mathbf{r})+f(\theta,\phi)\exp(ikr)/r,$$

where the vector $\mathbf{k}_i$ is in the direction of the incident plane wave. The noncentral potential $U(\mathbf{r})$ may be represented by a Legendre expansion relative to some z axis

$$U(\mathbf{r})=\sum_{L\text{ even}} U_L(r)P_L(\cos\theta), \tag{3.1}$$

where the U_L are assumed to have the following asymptotic dependence as $r\to\infty$:

$$\begin{aligned} U_2&\sim q/r^3-p'/r^4+O(r^{-5}),\\ U_0&\sim -p/r^4+O(r^{-5}), \end{aligned} \tag{3.2}$$

and higher U_L, if they exist, are $O(r^{-5})$. The reason for the particular asymptotic dependence assumed for the potentials by Eq. (3.2) is that this is the asymptotic form which one finds for the electrostatic interaction of a charged particle with a neutral quantum mechanical system, for example an electron and an atom. In this case the asymptotic form of U_2 will be proportional to the electric quadrupole moment of the target system, and this will be the dominant long-range interaction if it exists. The r^{-4} term in U_0 will be proportional to the target's electrical polarizability. The derivation of the known asymptotic forms Eq. (3.2) for an electron-atom system is indicated in Sec. 4. The restriction of the sum to even terms only is also dictated by the application. Physically this follows from the target system's having a well-defined parity.[7] In addition to being noncentral, we may allow the potential $U(\mathbf{r})$ to be nonlocal and energy dependent, provided that the nonlocality is of short range and the energy dependence can be represented by an expansion in k^2 near $k=0$. (These points will not be explicitly discussed further.)

In order to apply the method discussed in the preceding section, the Schrödinger equation must first be

[7] The only exception to this is the excited states of the H atom, where there is degeneracy between states of different angular momentum.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

replaced by the appropriate first-order nonlinear equation for the quantity of interest, in this case the scattering amplitude, $f(\theta,\phi)$. This equation was found by Spruch[8] to be

$$f_r(\mathbf{k}_i\to\mathbf{k}_f)=f_{r_0}(\mathbf{k}_i\to\mathbf{k}_f)-(1/4\pi)\int_{r_0}^{r}r'^2dr'$$
$$\times\int d\Omega[e^{i(\mathbf{k}_i\cdot\mathbf{r}')}+f_{r'}(\mathbf{k}_i\to\mathbf{k}')e^{ikr'}/r']U(\mathbf{r}')$$
$$\times[e^{-i(\mathbf{k}_f\cdot\mathbf{r}')}+f_{r'}(-\mathbf{k}_f\to\mathbf{k}')e^{ikr'}/r'], \quad (3.3)$$

where the argument r is now written as a subscript. The vector $\mathbf{k}_i$ is again in the direction of the incident plane wave and $\mathbf{k}_f$ is a vector in the final direction, or the direction of observation. The vector $\mathbf{k}'$ has as its angles the variables θ and ϕ of integration. The lower limit of integration, r_0 is arbitrary.

Before looking for the contribution to f from the asymptotic region of the potential, let us first, as in the last section, consider the expansion of f_r for r less than some finite and fixed, though possibly large, R. It is easily verified that the appropriate expansion which takes the place of Eq. (2.2) of the last section is one in powers of ik, namely

$$f=\sum_{n=0}^{\infty}(ik)^n f_n \quad (3.4)$$

for $r<R$ fixed. Substituting this in Eq. (3.3) and equating like powers of ik one finds the equations for the individual functions, f_n. In particular the equation for f_0 is

$$f_0(r)=-(1/4\pi)\int_0^r r'^2dr'[1+f_0/r]^2\int d\Omega U(\mathbf{r}), \quad (3.5)$$

where for the present purpose we have set $r_0=0$. Since the angular integral over the potential simply gives $4\pi U_0$, we see that Eq. (3.5) is identical with Eq. (2.3) for the scattering length (rather for $-A$), with the potential given by the spherically symmetric component U_0. Therefore we shall refer to the quantity $-f_0$ as the scattering length and use the notation $-f_0=A$. An important property of A is that it is entirely independent of the $L>0$ (the nonspherically symmetric) terms in the potential.

The equation for the next function, f_1, in Eq. (3.4) may similarly be found. It is more complicated but simplifies when we use the fact that the potential $U(\mathbf{r})$ contains only even angular terms. We shall not write the equation for f_1, but it is exactly satisfied by the function

$$f_1(r)=f_0^2(r). \quad (3.6)$$

If we temporarily imagine that the potentials are of short range so that we may go to the limit $r=\infty$, Eq. (3.6) gives the optical theorem in the zero-energy limit. Again we note that Eq. (3.6) explicitly assumes that the potential has no component proportional to $P_1(\cos\theta)$.

We may now sum up the result of Eqs. (3.4) through (3.6) to write the expansion of the amplitude, f_R for any finite though possibly very large R, as $kR\to 0$. It is

$$f_R=-A(R)+ikA^2(R)+O(k^2). \quad (3.7)$$

Equation (3.7) might have been simply written down immediately as the known contribution to the scattering amplitude due to a potential in the region r less than some finite R. However, it seems more consistent to rederive it in the context of the present method.

We now proceed to our main purpose, namely to seek the contribution to the amplitude f from the region $r>R$, using Eq. (3.3) and writing $r_0=R$. As in the last section, the arbitrary distance R will be considered to be very large so that the asymptotic expansion (3.2) of the potential may be used. Accordingly we substitute the potential from (3.1) and (3.2) into Eq. (3.3) and multiply out the two brackets. We again introduce new variables

$$x=kr',$$
$$y=Kr'=(K/k)x, \quad (3.8)$$

where K is the magnitude of the momentum transfer vector, $\mathbf{K}=\mathbf{k}_i-\mathbf{k}_f$. It is related to k by $K=2k\sin(\theta_{if}/2)$ (see Fig. 1). The equation becomes

$$f_r(\mathbf{k}_i\to\mathbf{k}_f)=f_R(\mathbf{k}_i\to\mathbf{k}_f)-(1/4\pi)\int_{kR}^{kr}x^2dx$$
$$\times\int d\Omega\{qP_2(\cos\theta)x^{-3}-kx^{-4}[p+p'P_2(\cos\theta)]+\cdots\}$$
$$\times\{e^{iy\cos(\theta,\theta_K)}+(k/x)f_{r'}(\mathbf{k}_i\to\mathbf{k}')e^{ix[1-\cos(\theta_f,\theta)]}$$
$$+(k/x)f_{r'}(-\mathbf{k}_f\to\mathbf{k}')e^{ix[1+\cos(\theta,\theta_i)]}$$
$$+(k^2/x^2)f_{r'}(\mathbf{k}_i\to\mathbf{k}')f_{r'}(-\mathbf{k}_f\to\mathbf{k}')e^{2ix}\}, \quad (3.9)$$

where the notation $\cos(\theta_1,\theta_2)$ is used here and throughout the paper to denote the cosine of the angle, θ_{12}, between the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$. $\cos(\theta_1,\theta_2)$ is given by the relation

$$\cos(\theta_1,\theta_2)=\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2). \quad (3.10)$$

There remains nothing now but to go ahead and evaluate the integrals in Eq. (3.9) by the iteration or perturbation method, beginning with

$$f_r^{(0)}=f_R=-A(R)+ikA^2(R)+O(k^2), \quad (3.11)$$

where the superscript henceforth denotes the order of perturbation theory. We proceed to sufficiently high order in perturbation theory to obtain all the powers of k which are desired. As implied in Eq. (3.11) only

[8] Reference 6, p. 209.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

terms through $O(k)$ will be sought here. (These will be seen to provide more than sufficient difficulty in evaluating while at the same time furnishing a great deal of detail in the results.)

By inspection of Eq. (3.9) it may be seen directly to what order in perturbation theory we must go in order to find all terms through $O(k)$, by taking the heuristic point of view explained in the last section of neglecting temporarily the k dependence of the lower limit of integration. Of the four exponential terms in the last bracket, the first term is $O(1)$. This will be called the Born term since it appears also in the Born approximation. Since it is not multiplied by any factor f as are the subsequent terms, it makes its contribution once and for all in first order. It contributes to f a q term of $O(1)$ and p and p' terms of $O(k)$. Higher terms in the potential would give contributions of $O(k^2)$ which is why they have not been written explicitly. Skipping ahead to the last exponential term in the bracket, it is seen that this already contains a factor k^2 which makes it higher order from the very beginning, and so it makes no contribution. Finally the two linear f terms in the bracket behave alike. They have a factor of k to start with. Since the lowest order terms in $f^{(0)}$ and $f^{(1)}$ are $O(1)$, the p and p' terms produced by these will always be $O(k^2)$. In first order, these linear f terms will produce q terms of $O(k)$ proportional to $f^{(0)}$. When we set these f's equal to $f^{(1)}$ to obtain the second-order result, we see that the Born term proportional to q will result in a q^2k term in this order. If the iteration process is continued further, it is seen that only terms of $O(k^2)$ are produced and so second order suffices.

We now proceed to evaluate the first-order contribution, $f^{(1)}$ to the amplitude. First, the zero-order quantity, $f^{(0)}$, given by Eq. (3.11) is substituted for f on the right-hand side of Eq. (3.9). The angular integrations are easily performed using the relation

$$\int d\Omega P_L(\cos\theta)e^{ix\cos(\theta,\theta_k)}=4\pi i^L j_L(x)P_L(\cos\theta_k). \quad (3.12)$$

This reduces Eq. (3.9) to

$$\begin{aligned} f_r^{(1)}(\mathbf{k}_i\to\mathbf{k}_f)&-[-A(R)+ikA^2(R)] \\ &=P_2(\cos\theta_K)q\int_0^{kr} dx x^{-1}j_2(y)+pk\int_{kR}^{kr} dx x^{-2}j_0(y) \\ &\quad -p'kP_2(\cos\theta_K)\int_0^{kr} dx x^{-2}j_2(y) \\ &\quad -qkA(R)[P_2(\cos\theta_f)+P_2(\cos\theta_i)] \\ &\quad \times\int_0^{kr} dx x^{-2}e^{ix}j_2(x)+O(k^2). \quad (3.13) \end{aligned}$$

The lower limit of integration has been extended to zero whenever the error introduced into f thereby was $O(k^2)$. The first three terms, resulting from the Born term in Eq. (3.9) are complicated by the presence of two different variables, x and y. It is easiest to eliminate the variable x in favor of y by use of Eq. (3.8). In the first integral, this merely changes the limits ($kr\to Kr$). In the second and third it also replaces the initial factor k by K.

Before proceeding to second order, let us look at the physical amplitude, $f_\infty^{(1)}$ in this order. Letting $r\to\infty$, we have

$$\begin{aligned} f_\infty^{(1)}(\mathbf{k}_i\to\mathbf{k}_f)&=-[A(R)-p/R]+\tfrac{1}{3}qP_2(\cos\theta_K) \\ &\quad -\tfrac{1}{4}\pi pK-(\pi/16)p'KP_2(\cos\theta_K)+ik\{A^2(R) \\ &\quad -\tfrac{1}{12}A(R)q[P_2(\cos\theta_f)+P_2(\cos\theta_i)]\}+O(k^2). \quad (3.14) \end{aligned}$$

The first term on the right is of course the scattering length, partially renormalized in first order. The second term represents the principle new result of the paper, a term generated by the quadrupole potential which is $O(1)$ but with angular dependence. As in previous work, this leading term is given by the Born integral. The angular factor $P_2(\cos\theta_K)$ requires some explanation. The angle θ_K is not the scattering angle. It is the angle made with the z axis by the momentum transfer vector $\mathbf{K}=\mathbf{k}_i-\mathbf{k}_f$ (the magnitude, K, of this vector appears in the following term). The vector $\mathbf{K}$ is represented in Fig. 1. The magnitude K is given (since $|\mathbf{k}_i|=|\mathbf{k}_f|=k$) by the familiar relation

$$K=2k\sin(\theta_{if}/2)=k\sqrt{2}[1-\cos(\theta_i,\theta_f)]^{1/2}.$$

The quantity $\cos\theta_K$ can be found from the relation

$$\cos(\theta_K)=\frac{K_z}{K}=\frac{\cos\theta_i-\cos\theta_f}{\sqrt{2}[1-\cos(\theta_i,\theta_f)]^{1/2}}. \quad (3.15)$$

The denominator might in turn be eliminated in favor of the angles θ_i, θ_f and $\phi_i-\phi_f$ by the relation

$$\cos(\theta_i,\theta_f)=\cos\theta_i\cos\theta_f+\sin\theta_i\sin\theta_f\cos(\phi_i-\phi_f).$$

So if the initial angle θ_i is arbitrary relative to the z axis from which the quadrupole potential is measured, the amplitude will depend on ϕ as well as θ. Now for the special case $\theta_i=0$, Eq. (3.15) reduces to

$$\cos\theta_K=\sin(\theta/2)\quad(\theta_i=0),$$

where we have written $\theta_f=\theta$. $P_2(\cos\theta_K)$ is then given by

$$P_2(\cos\theta_K)=P_2[\sin(\theta/2)]=\tfrac{1}{4}(1-3\cos\theta)\quad(\theta_i=0). \quad (3.16)$$

The third term in Eq. (3.14) is the known term[4] giving the lowest order effect of the r^{-4} potential. As a consistency check, if one considers the case $\theta_i=0$, the imaginary terms for $\theta_f=0$ may be deduced from the two leading (real) terms by means of the optical theorem[9] and they agree with those in Eq. (3.14).

[9] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), p. 105.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Now, to proceed to second order, we substitute $f_r^{(1)}$ from Eq. (3.13) with the required changes of angles back into the right-hand side of Eq. (3.9). The leading term in $f_r^{(1)}-f^{(0)}$ is the first q term in (3.13) which is $O(1)$. This term multiplied with the q term in the potential will given a q^2k term as mentioned above. All other terms are $O(k^2)$ and hence higher order (except for second-order corrections to the scattering length arising from the p terms). Since further iterations may be seen to produce only higher order terms, $f^{(2)}$ is as high in perturbation theory as we shall go, and so, letting $r\to\infty$, we find

$$f_\infty^{(2)}(\mathbf{k}_i\to\mathbf{k}_f)-f_\infty^{(1)}=-(1/4\pi)\int_{kR}^{\infty}x^2dx \times\int d\Omega(q/x^3)P_2(\cos\theta)(k/x)qP_2(\cos\theta_{K'}) \times[\tfrac{1}{3}-j_1(K'r')/K'r']e^{ix[1-\cos(\theta_f,\theta)]} +(\mathbf{k}_i\leftrightarrow-\mathbf{k}_f)+O(k^2), \quad (3.17)$$

where, according to the requirements of Eq. (3.9), $\theta_{K'}$ is defined as the angle made with the z axis by the vector $\mathbf{K}'=\mathbf{k}_i-\mathbf{k}'$. An analysis like that leading to Eq. (3.15) for θ_K gives

$$\cos\theta_{K'}=\frac{\cos\theta_i-\cos\theta}{\sqrt{2}[1-\cos(\theta_i,\theta)]^{1/2}}, \quad (3.18)$$

while the magnitude K' is equal to $k\sqrt{2}[1-\cos(\theta_i,\theta)]^{1/2}$. The corresponding quantities for $(\mathbf{k}_i\leftrightarrow-\mathbf{k}_f)$ should be clear. The r' should be eliminated by the relation $K'r'=(K'/k)x$. At this point it does not seem that the integrations in Eq. (3.17) can be carried out analytically, so we will simply write

$$f_\infty^{(2)}-f_\infty^{(1)}=q^2kI(\theta,\phi)+O(k^2), \quad (3.19)$$

where

$$I(\theta,\phi)=-(1/4\pi)\int_0^\infty dx x^{-2}e^{ix} \times\int d\Omega P_2(\cos\theta)P_2(\cos\theta_{K'}) \times[\tfrac{1}{3}-j_1(K'r)/K'r]e^{ix\cos(\theta_f,\theta)} +(k_i\leftrightarrow-k_f)$$

$$=(4\pi)^{-2}\int d\tau_x x^{-4}e^{ix}e^{-ix\cos(\theta_f,\theta)}P_2(\cos\theta) \times\int_0^x x'^2dx'\int d\Omega' x'^{-3}P_2(\cos\theta') \times e^{ix'\cos(\theta',\theta_{K''})}+(\mathbf{k}_i\leftrightarrow-\mathbf{k}_f). \quad (3.20)$$

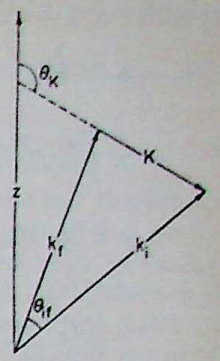


FIG. 1. Shows schematically the meaning of the angle θ_K.

$I(\theta,\phi)$ has been written in the alternate form of a double integral from which it arose, since it might prove more amenable to solution in this form. At any event it should be possible to evaluate the integrals numerically for various values of θ_f and ϕ_f if it is ever needed. The angle $\theta_{K''}$ represents the vector $\mathbf{K}''$ which is the difference between the vectors $\mathbf{k}_i$ and $\mathbf{k}'$, the latter having the angles θ' and ϕ' of integration. $I(\theta,\phi)$ should of course have both a real and an imaginary part.

As mentioned earlier, an evaluation of $f^{(3)}$ leads only to terms of $O(k^2)$ aside from further renormalizations of the scattering length, A. The final result then may be written, combining Eqs. (3.19) and (3.14),

$$f(\mathbf{k}_i\to\mathbf{k}_f)=-A+\tfrac{1}{3}qP_2(\cos\theta_K) -\tfrac{1}{4}\pi pK+q^2k\,\mathrm{Re}I(\theta,\phi)-(\pi/16)p'KP_2(\cos\theta_K) +ik[A^2-\tfrac{1}{12}qA(P_2(\cos\theta_i)+P_2(\cos\theta_f)) +q^2\,\mathrm{Im}I(\theta,\phi)]+O(k^2), \quad (3.21)$$

where the renormalized quantity, $A=A(\infty)$ has been substituted throughout, and θ_K may be found from Eq. (3.15), ff. Eq. (3.21), the final result of this section represents the expansion of the scattering amplitude through terms linear in k for the potential of Eqs. (3.1) and (3.2). An attempt will be made in Sec. 4 to apply this result to low-energy electron-atom scattering.

Dipole Potential

As a digression, let us attempt to apply the above method to a potential whose asymptotic form contains a leading dipole term

$$U(\mathbf{r})\sim D\cos(\theta)/r^2. \quad (3.22)$$

(This kind of potential arises in connection with the scattering of electrons from the excited states of the H atom, because of the accidental degeneracy between states of different angular momentum.) We substitute this asymptotic form into Eq. (3.3), neglecting any higher order terms in U since we shall only look for the leading k term in $f(\theta)$. In addition, we again substitute the new variables $kr'=x$ and $Kr'=y$, and multiply through by a factor k. We find in place of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Eq. (3.9)

$$kf_r(\mathbf{k}_i \to \mathbf{k}_f) = kf_R - (1/4\pi)\int_{kR}^{kr} x^2 dx \times \int d\Omega D\cos(\theta) x^{-2}\left[e^{iy\cos(\theta,\theta_K)} + x^{-1}kf_x(\mathbf{k}_i \to \mathbf{k}')e^{ix}e^{-ix\cos(\theta,\theta_f)} + x^{-1}kf_x(-\mathbf{k}_f \to \mathbf{k}')e^{ix}e^{ix\cos(\theta,\theta_i)} + x^{-2}kf_x(\mathbf{k}_i \to \mathbf{k}')kf_x(-\mathbf{k}_f \to \mathbf{k}')e^{2ix}\right]. \quad (3.23)$$

Now by Eq. (3.11) f_R is $O(1)$. (As mentioned before the ik term would be different for a dipole potential, but this is a higher order term anyway.) So kf_R is $O(k)$. Let us neglect terms of $O(k)$ as higher order and look only for terms of lower order than this. We then have $kf^{(0)} = O(k)$. Using this on the right-hand side we find for $f^{(1)}$

$$kf^{(1)}(\mathbf{k}_i \to \mathbf{k}_f) = -\left(\frac{D}{4\pi}\right)\int_{kR}^{kr} dx \times \int d\Omega \cos\theta e^{iy\cos(\theta,\theta_K)} + O(k) = -iD(k/K)\cos\theta_K[1 - j_0(y)] + O(k). \quad (3.24)$$

Thus, in first order, kf is $O(1)$ as we could see by inspection of (3.23). If we now proceed to substitute $kf^{(1)}$ on the right-hand side in order to find $kf^{(2)}$, we see immediately that the new terms are also $O(1)$ and so to all higher orders of perturbation theory. So that the situation here is very different from that for an r^{-3} or r^{-4} potential, where the new terms in higher orders of perturbation theory represented successively higher powers of k. With the dipole potential, all orders of perturbation theory give the same power of k, namely $f \propto k^{-1}$. Thus for the dipole potential, Eq. (3.24) (for $y = \infty$) gives only the first Born approximation to the k^{-1} term in $f(\theta)$, and so will only be of use if D happens to be very small. For moderate D higher order perturbation terms would be needed, while for still higher D the expansion will very likely not converge at all.

So we see that unfortunately the present method breaks down for a potential which falls off as slowly as r^{-2}. From Eq. (3.24) we may at least salvage the qualitative result that the amplitude diverges as k^{-1} in agreement with the classical result,[10] and also with that of Seaton[11] who considered the electron-hydrogen problem for a fixed value of the total angular momentum L, and with Gailitis and Damburg[12] who carried this result somewhat farther. If one wishes to go beyond the Born approximation or the contribution to the amplitude from single angular momentum states, what seems to be needed is an exact solution of the wave equation with the dipole potential (3.22).

4. ELECTRON-ATOM SCATTERING

The particular asymptotic form of the potential $U(r)$ given by Eq. (1.1) was chosen because it corresponds to the long-range interaction between a neutral atom (or molecule) in a particular state and a charged particle (electron).

The potential between an atom and an electron is taken (neglecting spin-dependent forces) to be the sum of all Coulomb forces between particles (electrons and nucleus). When the separation r is very large, the interaction assumes the asymptotic form

$$V(\mathbf{r},\mathbf{r}_j) \sim \sum_{L=1}^{\infty} e^2/r^{(L+1)} \sum_{j=1}^{Z} r_j{}^L P_L[\cos(\theta,\theta_j)], \quad (4.1)$$

where Z is the atomic number of the atom, r and θ are the coordinates of the separated electron, and $\mathbf{r}_j$ represents the position of the jth atomic electron. The absence of an $L=0$ term results from the assumed neutrality of the atom. From this, the limiting form Eq. (1.1) of the effective potential seen by the scattered electron has been derived by a number of authors.[13]

The simplest way to derive an effective potential is to treat the asymptotic potential (4.1) as a perturbation of the atomic ground state and to find its expectation value in first- and second-order perturbation theory,[14] treating r as a (large) parameter. If we assume the atom to be and to remain in a particular state of magnetic quantum number M, the first-order result is

$$V_{00} \sim -\frac{e}{2}\sum_{L=1}^{\infty} \frac{Q_L(M)P_L(\cos\theta)}{r^{L+1}}, \quad (4.2)$$

where the $Q_L(M)$ are the permanent electric multipole moments of the atom in its ground state with quantum numbers J and M. Only those moments will exist for which $L \leq 2J$, so that for example a spherically symmetric atom will have no permanent moments. Further, since all atoms in their ground states have a well-defined parity,[15] all the odd moments vanish so that the leading term in (4.2) is that for $L=2$, the quadrupole term. This immediately gives us the first term of Eq. (1.1). The M dependence of the quadrupole

[10] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., Cambridge, Massachusetts, 1953), p. 91.
[11] M. J. Seaton, Proc. Phys. Soc. (London) **77**, 174 (1961).
[12] M. Gailitis and R. Damburg, Proc. Phys. Soc. (London) **82**, 192 (1963).
[13] See for example J. Holtzmark, Z. Physik **55**, 437 (1929); E. Gerjuoy and S. Stein, Phys. Rev. **97**, 1671 (1955).
[14] Reference 9, p. 152, ff.
[15] This is generally true of excited states as well with hydrogen being an important exception.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

moment is given by[16]

$$Q(M)=\frac{3M^2-J(J+1)}{J(2J-1)}Q, \tag{4.3}$$

where the number Q is defined[17] as the expectation value of the quadrupole operator for the state $M=J$,

$$\tfrac{1}{2}Q=e\int d\tau \sum_{j=1}^{Z} r_j^2 P_2(\cos\theta_j)\,|\psi_0|^2_{M=J} \tag{4.4}$$

with the integral extending over all the coordinates of the atom (a sum over spins is implied). Equation (4.3) has as a consequence that the average of $Q(M)$ over M vanishes

$$Q_{\rm av}=0. \tag{4.5}$$

Unfortunately the physical situation is somewhat more complicated. Due to the degeneracy of states with the same magnetic quantum number M, *transitions* between these states may be caused by electrons with vanishingly small energy, so that the first-order potential should really be written as a matrix of dimension $2J+1$. What has been done in (4.2) is to neglect all but "elastic" collisions, i.e., those in which M remains constant. This neglect is not justified, and an attempt is made below to remedy it.

Returning to Eq. (4.1) and evaluating it now in second-order perturbation theory gives

$$V^{(2)}\sim -e^2/(2r^4)[\alpha+\alpha'(M)P_2(\cos\theta)]+O(r^{-6}), \tag{4.6}$$

where α is the (average) electric polarizability of the atom and α' (which exists only for nonsymmetric atoms) has the property that, like the quadrupole moment, its average over M vanishes. In evaluating this result, the variable r is treated as a constant parameter. This is known as the adiabatic approximation. However, it has been shown[18] that the result gives the correct asymptotic behavior of the effective potential, in other words that the adiabatic approximation is correct asymptotically. This gives rise to the second term in Eq. (1.1).

A more consistent and satisfying way of arriving at the long-range electron-atom potential is to start from the exact one-body optical potential equation[19] for an electron scattered from an atom in its ground state

$$[\nabla^2+k^2-(2m/\hbar^2)V_{\rm opt}]u(\mathbf{r})=0, \tag{4.7}$$

where

$$V_{\rm opt}(\mathbf{r})=V_{00}+\mathbf{V}(E-\mathbf{H})^{-1}\mathbf{V}^{+}. \tag{4.8}$$

V_{00} is the expectation value of the full electron-atom potential averaged over the target ground state. It corresponds exactly with the first-order perturbation theory result whose asymptotic form is given in Eq. (4.2) and the same comments made there again apply. In the second term, the vector $\mathbf{V}$ has as its components the matrix elements V_{0i} connecting the target ground state with each of its excited states. Similarly the matrix $\mathbf{H}$ has components H_{ij} connecting any 2 excited states of the atom (but with no ground-state components). The Pauli principle may be considered to be implicit in (4.8) or else exchange terms may be added explicitly. In either case it may easily be shown[20] that its effects fall off exponentially as $r\to\infty$ and so will not affect the asymptotic behavior of $V_{\rm opt}$.

Now if one lets $r\to\infty$ in Eq. (4.8) using (4.1) for the potential one obtains in a straight forward manner Eq. (4.6) as the limiting form for the second term but with error this time[20] of $O(r^{-5}d/dr)$. And so the asymptotic form of the optical potential (with the restriction mentioned below to "elastic," i.e., M conserving, collisions) is

$$\frac{2m}{\hbar^2}V_{\rm opt}\sim\frac{Q(M)}{ea_0}\frac{P_2(\cos\theta)}{r^3}-\frac{1}{a_0r^4}[\alpha+\alpha'(M)P_2(\cos\theta)]+O\left(\frac{1}{r^5}\right), \tag{4.9}$$

where $a_0=\hbar^2/me^2$ is the electron Bohr radius. Comparing Eq. (4.9) with Eq. (1.1) we see that in applying the formulas of Sec. 3 to this case we should set

$$q=Q(M)/ea_0, \quad p=\alpha/a_0, \quad p'=\alpha'(M)/a_0. \tag{4.10}$$

The foregoing, however is not really adequate to describe the scattering of electrons by atoms, since transitions of the atom from one magnetic state to another have not been taken into account.

Allowance for Transitions Between Magnetic States

If the target atom is initially in a magnetic state M relative to some z axis, and an electron beam is incident with momentum $\mathbf{k}_i$, the asymptotic form of the wave function may be written

$$\psi\sim\psi_M(\mathbf{r}_j)e^{i\mathbf{k}_i\cdot\mathbf{r}}+\sum_{M'=-J}^{J} f_{M,M'}(\theta,\phi)\psi_{M'}(\mathbf{r}_j)e^{ikr}/r, \tag{4.11}$$

where $\psi_{M'}(\mathbf{r}_j)$ are the ground-state wave functions for

[16] J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley & Sons, New York, 1952), p. 28.

[17] We have used the definition of Blatt and Weisskopf. The molecular definition of the quadrupole moment used by Gerjuoy and Stein (Ref. 13) can be made to agree with this by the relation, $Q=-(2J/2J+3)Q_{GS}$.

[18] L. Castillejo, I. C. Percival, and M. J. Seaton, Proc. Roy. Soc. (London) A254, 259 (1960).

[19] H. Feshback, Ann. Phys. (N. Y.) 5, 357 (1958). See also M. H. Mittleman and K. M. Watson, Phys. Rev. 113, 198 (1959); B. A. Lippmann, M. H. Mittleman, and K. M. Watson, *ibid.* 116, 920 (1959).

[20] M. H. Mittleman, Ann. Phys. (N. Y.) 14, 94 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the atom in the magnetic state M'. The amplitude $f(\theta,\phi)$ is now a matrix of dimension $2J+1$, connecting these states. Correspondingly the potential V_{00} of Eq. (4.8) is likewise a matrix, as mentioned earlier. If the matrix elements of the potential V_{00} connecting the states M and M' are labeled $V_{00}(M,M')$, where V is still given by Eq. (4.1) the asymptotic form of the potential matrix elements is given by

$$(2m/\hbar^2)V_{00}(M,M')\sim r^{-3}(4\pi/5)^{1/2}Y_{2,M-M'}(\theta,\phi)\times(Q_{M,M'}/ea_0)+O(r^{-5}), \quad (4.12)$$

where

$$\tfrac{1}{2}Q_{M,M'}=e(4\pi/5)^{1/2}\int d\tau\psi_{M'}{}^*\times\sum_{j=1}^{Z} r_j^2 Y_{2,M-M'}{}^*(\theta_j,\phi_j)\psi_M. \quad (4.13)$$

The $Q_{M,M'}$ are related to the ordinary quadrupole moment, $Q=Q_{J,J}$ by the relation

$$Q_{M,M'}=\left[\frac{(J+1)(2J+3)}{J(2J-1)}\right]^{1/2}\times C(J,2,J;M',M-M')Q, \quad (4.14)$$

where Q is the quadrupole moment defined by Eq. (4.4), and C is the Clebsch Gordon coefficient.[21] The $Q(M)$ defined before are simply the diagonal elements, Q_{MM}.

The second term of the optical potential, Eq. (4.8), is still $O(r^{-4})$, so that Eq. (4.12) represents the leading term in the optical potential matrix as $r\to\infty$. If the amplitudes $f_{M,M'}$ are now written as functions of r as was done in Sec. 3 for f, they satisfy coupled equations. These equations may be derived from the coupled optical potential equations for the coefficients of the various ψ_M, whose asymptotic forms appear in Eq. (4.11), exactly as was done for the amplitude f in Ref. 8. The equations are

$$f_{M,M';r}(\mathbf{k}_i\to\mathbf{k}_f)=-(1/4\pi)\int_0^r r'^2dr'\int d\Omega'\times\sum_{M''}\sum_{M'''}\psi_{M''}{}^{M}(\mathbf{r}';\mathbf{k}_i\to\mathbf{k}')U_{M'',M'''}(\mathbf{r}')\times\psi_{M'''}{}^{M'}(\mathbf{r}';-\mathbf{k}_f\to\mathbf{k}'),$$

where

$$\psi_{M''}{}^{M}(\mathbf{r}';\mathbf{k}_j\to\mathbf{k}')=\delta_{M,M''}e^{i\mathbf{k}_j\cdot\mathbf{r}'}+f_{M,M'';r'}(\mathbf{k}_j\to\mathbf{k}')e^{ikr'}/r'.$$

The vector $\mathbf{k}'$ again has the angles Ω' of integration. The sums are over all magnetic substates, and the asymptotic form of the $U_{M'',M'''}$ are given by Eq. (4.12) plus terms of $O(r^{-4})$. If we are satisfied with terms in the amplitude of $O(1)$, we see that these were given in Sec. 3 by the Born term in Eq. (3.3), together with the r^{-3} term in the potential, apart from contributions from the short-range part of the potential. The same will be true in the present case. The Born term for the $f_{M,M'}$, with the potential given by Eq. (4.12), is

$$f_{M,M'}(\mathbf{k}_i\to\mathbf{k}_f)=f_{M,M';R}-\frac{1}{4\pi}\int_R^\infty r^2dr\int d\Omega e^{i\mathbf{K}\cdot\mathbf{r}}\times\frac{1}{r^3}\left(\frac{4\pi}{5}\right)^{1/2}Y_{2,M-M'}(\theta,\phi)\times\frac{Q_{M,M'}}{ea_0}+O(k)$$

$$=-A\delta_{M,M'}+\frac{1}{3}\frac{Q_{M,M'}}{ea_0}\left(\frac{4\pi}{5}\right)^{1/2}\times Y_{2,M-M'}(\theta_K,\phi_K)+O(k). \quad (4.15)$$

$\mathbf{K}$ is again the momentum transfer vector $\mathbf{k}_i-\mathbf{k}_f$, A has the same meaning as in Sec. 3, and the substitution of A for $A(R)$ has again been made as in Secs. 2 and 3. The relation of θ_K to the angles of $\mathbf{k}_i$ and $\mathbf{k}_f$ was given by Eq. (3.15). Similarly $\tan\phi_K=K_x/K_y$ and this is equal to

$$\tan\phi_K=\frac{\sin\theta_i\sin\phi_i-\sin\theta_f\sin\phi_f}{\sin\theta_i\cos\phi_i-\sin\theta_f\cos\phi_f}. \quad (4.16)$$

The delta function arises from the fact that the (scattering length) equation which determines the leading term of f_R contains only the angular average of the potential and so does not connect states of different M.

Equation (4.15) represents the final result of this section, the scattering amplitude to lowest order in k for collisions of very slow electrons with a nonsymmetric atom initially in magnetic state M. We shall now mention two obvious applications of this formula. The first is the scattering of slow electrons from unoriented atoms. For this case one should construct the cross sections and then average over initial states M and sum over final states M'. Using Eq. (4.14) for $Q_{M,M'}$ we find

$$\sigma(\theta)_{\rm av}=\frac{1}{2J+1}\sum_M\sum_{M'}|f_{M,M'}|^2=A^2+\frac{1}{9}\left(\frac{Q}{ea_0}\right)^2\times\frac{(J+1)(2J+3)}{5J(2J-1)}+O(k), \quad (4.17)$$

where J is the total angular momentum of the atom.[22]

[21] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957), p. 33 ff.; E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (The University Press, Cambridge, 1951), p. 76 ff.

[22] It is interesting to compare the above result with those of Gerjuoy and Stein, Ref. 13, although they were not concerned with elastic scattering. If one nevertheless sets $J_b=J_a$ in their results and $k_b=k_a$ and uses the relation between the quadrupole moments mentioned in Ref. 17, one reproduces the second term of Eq. (4.17). (There is of course no A^2 term in these results since they did not consider elastic scattering.)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We have used the sum rules for spherical harmonics and for Clebsch Gordan coefficients[23] together with Eq. (4.5). It should be noted that the differential cross section given by Eq. (4.17) is spherically symmetric.

If the target atoms have been *oriented* so that they are initially in magnetic state M, then Eq. (4.15) may be used directly to determine the differential cross section for transitions to any other state M' caused by collisions with very slow electrons. In this case, changes of M by 0, 1, and 2 units are allowed. To find the total cross sections for these transitions, the integrals over ϕ_f and θ_f may be done, after substituting Eq. (3.15) for θ_K, giving

$$\sigma_{M,M}=4\pi A^2-(2/3)\pi A\frac{Q_{M,M}}{ea_0}P_2(\cos\theta_i)+(1/9)\pi\left(\frac{Q_{MM}}{ea_0}\right)^2[1-(3/8)(\sin\theta_i)^4]+O(k),$$

$$\sigma_{M,M\pm1}=(1/12)\pi\left(\frac{Q_{M,M\pm1}}{ea_0}\right)^2[1+(1/3)(\cos\theta_i)^4]+O(k), \quad (4.18)$$

$$\sigma_{M,M\pm2}=(5/48)\pi\left(\frac{Q_{M,M\pm2}}{ea_0}\right)^2[1-(2/5)(\cos\theta_i)^2-(1/15)(\cos\theta_i)^4]+O(k).$$

These total cross sections still depend on the angle θ_i between the direction of the incident electron and the z axis along which the atoms have been oriented.

Finally with regard to the energy range of validity of Eq. (4.15) and the results of Sec. 3, as in previous work[4] necessary conditions can be found for the energy. In the iteration process begun in Sec. 3 and implied in the work of the present section, the iterated quantities consisted of the terms qk and pk^2 each multiplied by a factor of order unity. Necessary conditions then on the applicability of the preceding results is that dimensionless quantities kQ/ea_0 and $k^2\alpha/a_0$ be small compared to unity. The latter condition is that already found for the polarization potential.

ACKNOWLEDGMENTS

The author would like to thank Dr. Lewis M. Branscomb and the Fellows and staff of the Joint Institute for Laboratory Astrophysics for their hospitality during his stay at the Institute.

[23] See Ref. 21 (Rose), pp. 60 and 34.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Lifetime of the First $^1P_1$ State of Zinc, Calcium, and Strontium*

ALLEN LURIO
IBM Watson Laboratory, Columbia University, New York, New York

AND

R. L. DEZAFRA† AND ROBERT J. GOSHEN
Columbia Radiation Laboratory, Columbia University, New York, New York

(Received 9 January 1964)

The lifetime of the first excited $^1P_1$ state of zinc, calcium, and strontium has been measured by the zero-field level-crossing technique. The results are: zinc $\tau(^1P_1)=1.41\pm0.04\times10^{-9}$ sec, calcium $\tau(^1P_1)=4.48\pm0.15\times10^{-9}$ sec, and strontium $\tau(^1P_1)=4.97\pm0.15\times10^{-9}$ sec. The results are compared to earlier experimental results and also to theoretical predictions. It is found that our results are in excellent agreement with the Bates and Damgaard Coulomb approximation method for estimating the oscillator strengths.

I. INTRODUCTION

THE problem of determining atomic transition probabilities and excited-state lifetimes is an old one in physics, although it has received relatively little experimental attention in recent years. At the same time, advances in photoelectric recording and optical techniques make many early methods well worth reviewing and revising. Surveys of existing literature show a few advances in techniques since the early 1930's, while revealing upon occasion a rather considerable state of disagreement—not uncommonly by an order of magnitude or more—in many "absolute" measurements. Several excellent compilations of previous literature, tables of numerical values, and review articles exist. The best reference for all of these as well as individual articles, up to 1962, is the National Bureau of Standards monograph by Glennon and Wiese.[1]

Most absolute experimental determinations of oscillator strengths or transition probabilities of spectral lines have used the "hook" method developed by Rozhdestvenski and co-workers at Leningrad over the years from 1912 to the present.[2] This method and the various line absorption methods of measuring absolute oscillator strengths all require a knowledge of the density of the scattering atoms.[3] It is this requirement that presents a limitation to the precision of these methods. (For the measurement of *relative* oscillator strengths, however, the "hook" method seems to be excellent.) The zero-field level crossing or Hanle effect,[4] on the other hand, does not suffer from this requirement in the limit of low vapor densities, and should therefore be much more reliable. The level-crossing technique does require the production of strong unself-reversed resonance lines and consequently is limited to measurements of the oscillator strengths of the first few resonance lines of an element. In this respect the "hook" method would seem superior since it does not require resonance lines.

In a previous paper[5] the measurement of the lifetime of the first excited $^1P_1$ state of Cd by the zero-field level-crossing technique is reported. In the present paper we report on the lifetime of the first excited $^1P_1$ state of Ca, Sr, and Zn. In these experiments the resonance radiation is scattered from an atomic beam so that the difficulty of finding a transparent cell which is not attacked by Ca or Sr vapor is avoided. Our results are compared to previous measurements and to theoretical calculations for the oscillator strengths and are found to agree surprisingly well with the Bates and Damgaard approximation.[6]

II. THEORY

We first present a simple and somewhat restricted theory derived on strictly classical grounds, to give as directly as possible a physical picture of the Hanle or zero-field level-crossing effect. Later, a more generalized quantum-mechanical treatment is given leading to the same conclusions.

More than one geometrical arrangement is possible, and polarized light need not be used, but the simplest arrangement conceptually is probably the following. Consider a vapor of the atoms to be investigated—either in the form of a free atomic beam, or in a closed evacuated container—with the scattering region at the coordinate origin (see Fig. 1). A beam of resonance radiation with the **E** field polarized in the Y direction is sent in along the X axis. Absorption of the resonant light by the vapor raises the atoms to the excited state whose lifetime is to be measured. Classically one may consider the resonance light to excite electric dipoles

* Work supported in part by the U. S. Navy Office of Naval Research under Contract Nonr 266(45), in part by the National Aeronautics and Space Administration grant NsG-360, and in part by the Joint Services (U. S. Army, U. S. Navy Office of Naval Research, and U. S. Air Force Office of Scientific Research).

† Permanent address: State University of New York at Stony Brook, Stony Brook, Long Island, New York.

[1] B. M. Glennon and W. L. Wiese, Natl. Bur. Std. (U. S.), Monograph No. 50 (1962).

[2] See various articles in *Optical Transition Probabilities*, a collection of translated Russian articles published by the National Science Foundation, Washington, D. C., 1962.

[3] A. C. G. Mitchell and M. W. Zemansky, *Resonance Radiation and Excited Atoms* (Cambridge University Press, New York, 1961).

[4] W. Hanle, Z. Physik 30, 93 (1924); Ergeb. Exakt. Naturw. 4, 214 (1925).

[5] A. Lurio and R. Novick, Phys. Rev. 134, A608 (1964).

[6] D. R. Bates and A. Damgaard, Phil. Trans. 242, 101 (1949).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

oscillating along the direction of the Y axis.[7] These dipoles then re-radiate energy with the usual dipole radiation pattern, producing maximum intensity perpendicular to, and zero intensity along, the Y axis. However, if a magnetic field is present, aligned with the Z axis, the dipoles will precess about this field at their Larmor frequency ω until they re-radiate. ω depends on the field strength H via the Larmor relation $\omega=\gamma H$ where γ is the gyromagnetic ratio: in this case

$$\gamma=\omega/H=g_J\mu_0/\hbar, \qquad (1)$$

with g_J = the electronic g factor and μ_0 = the Bohr magneton. The angle of rotation through which a given dipole precesses before re-radiating is thus dependent on H. Multiplying the $\sin^2$ dipole-radiation pattern by an exponential damping factor to account for the random finite excited-state lifetimes over a large continuously excited collection of atoms, and integrating over all times, yields the following results for the observed intensity of scattered resonant radiation along the Y axis (at 90° to the incident beam) as a function of the magnetic field strength H:

$$I=C\int_0^\infty e^{-\Gamma t}\sin^2\omega t dt=\frac{C}{2}\left[1-\frac{\Gamma^2}{\Gamma^2+4\gamma^2H^2}\right], \qquad (2)$$

where Γ is the reciprocal of the mean life of the excited state and C is an arbitrary constant relating to the incident light intensity, the experimental geometry, and the number of scatterers. The result is just an inverted Lorentzian line shape as a function of H, with a full width at half-maximum equal to Γ. The half-maximum is seen to occur at a field value $H_{1/2}$ for which $4\gamma^2H^2_{1/2}=\Gamma^2$, or using Eq. 1, when

$$\Gamma=(2g_J\mu_0H_{1/2}/\hbar)\ \text{sec}^{-1}. \qquad (3)$$

Thus, a simple determination of the value of the magnetic field required to produce the half-maximum level in the Lorentzian line shape serves in principle to determine the mean life $\tau=1/\Gamma$, since g_J is generally well known. In practice fairly detailed profiles of the scattered intensity were taken as a function of H, as described in the next section, in order to carefully analyze and verify the Lorentzian shape. If the incident light along the X axis is unpolarized, one may carry out the theoretical analysis by resolving this light into two orthogonal polarizations. For the component polarized parallel to the Y axis the treatment is the same as given above. For the component polarized parallel to the Z axis, the scattered light is independent of the applied magnetic field since a dipole oscillating parallel to H does not precess. In practice one polarizes the incident light parallel to the Y axis in order to eliminate this constant background.

[7] The classical treatment gives the correct results for the simple case of the excitation from a $J=0$ ground state to a $J=1$ excited state. We have not examined the classical treatment in more complicated cases.

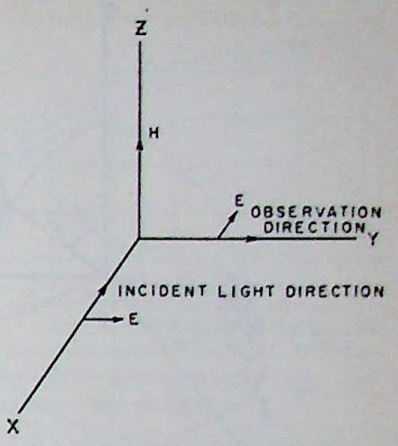


FIG. 1. Geometry used in Hanle effect experiment.

In odd A isotopes the presence of hyperfine structure may alter the Lorentzian profile in particular cases. This occurs since the g_F value of the odd isotopes will differ from the g_J value for the even isotopes. Moreover, the Hanle effect is only a special case at zero field of the more general "level-crossing" phenomenon discovered by Franken and co-workers[8] and the observational technique used here will also detect these nonzero-field crossings. This may then introduce distortions in the wings of the Lorentzian profile if fields large enough to reach nonzero-field level crossings are used, which may be the case with comparatively small hyperfine intervals and short lifetimes (e.g., $\tau\lesssim10^{-9}$ sec). This was not of importance for the elements studied here since the naturally occuring group II elements contain at most small amounts of odd A isotopes and those odd A isotopes which do occur all have $I\geq\frac{5}{2}$. Thus for these isotopes the level-crossing signals are a very small fraction of the scattered resonance radiation because of the large number of Zeeman levels.

We shall now give a quantum-mechanical derivation for the intensity of the scattered resonance radiation from a collection of identical atoms with $I=0$, excited from the $^1S_0$ ground state to the $^1P_1$ state and decaying back to the $^1S_0$ state. We will take the externally applied static magnetic field to be along the Z axis but will allow for arbitrary incident light and observing directions (see Fig. 2). The general expression for the rate of scattering of resonance radiation has been given by Breit[9] and Franken[8] and is

$$R(f,g)\propto N\sum_{\substack{\mu\mu'\\ mm'}}\frac{f_{\mu m}f_{m\mu'}g_{\mu' m'}g_{m'\mu}}{\Gamma-i[(E_\mu-E_{\mu'})/\hbar]}, \qquad (4)$$

where $f_{\mu m}=(\mu|\hat{e}\cdot\mathbf{r}|m)$, $g_{\mu m}=(\mu|\hat{e}'\cdot\mathbf{r}'|m)$, μ and μ' refer to the excited state magnetic sublevels, m and m' refer to the ground-state sublevels, and E_μ is the energy of the μth excited state. Figure 2 shows the spherical

[8] P. Franken, Phys. Rev. 121, 508 (1961).
[9] G. Breit, Rev. Mod. Phys. 5, 91 (1933).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

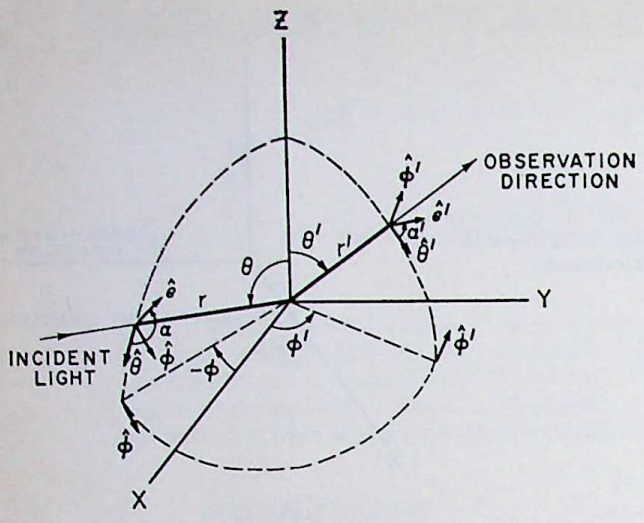


FIG. 2. Coordinate and polarization system used in quantum mechanical calculation of zero-field level-crossing signal.

coordinate system used to evaluate R. The incident light is directed toward the origin at coordinate angles θ and φ and the light polarization vector $\hat{e}$ makes an angle α with respect to $\boldsymbol{\theta}$ and lies in a plane normal to $\mathbf{r}$. $\hat{e}'$ and $\mathbf{r}'$ refer to the scattered light. With this coordinate system we have

$$\hat{e}=\hat{\imath}(\cos\theta\cos\varphi\cos\alpha-\sin\varphi\sin\alpha)+\hat{\jmath}(\cos\theta\sin\varphi\cos\alpha+\cos\varphi\sin\alpha)-\hat{k}(\sin\theta\cos\alpha). \quad (5)$$

We may define a new set of complex vectors which will prove useful by the relations

$$\begin{aligned} e_{\pm}&=e_x\pm ie_y=(\cos\theta\cos\alpha\pm i\sin\alpha)e^{\pm i\varphi},\\ e_z&=-\sin\theta\cos\alpha,\\ r_{\pm}&=r_x\pm ir_y, \end{aligned} \quad (6)$$

so that $\hat{e}\cdot\mathbf{r}=\frac{1}{2}(e_+r_-+e_-r_+)+e_zr_z$. In order to evaluate $f_{m\mu}$ and $g_{m\mu}$ we need the matrix elements[10]

$$\begin{aligned} \langle\gamma,0,0|e_{\pm}r_{\mp}|\gamma',1,\pm1\rangle&=\mp\sqrt{2}e_{\pm}\langle r\rangle,\\ \langle\gamma,0,0|e_zr_z|\gamma',1,0\rangle&=e_z\langle r\rangle, \end{aligned} \quad (7)$$

where $\langle r\rangle$ is the reduced matrix element of $\mathbf{r}$. Let us now define the excitation matrix $F_{\mu\mu'}=\sum_m f_{\mu m}f_{m\mu'}$ or more explicitly

$$F_{\mu\mu'}=\begin{vmatrix} \frac{1}{2}(e_+e_-) & -e_-e_z/\sqrt{2} & -\frac{1}{2}e_-^2\\ -e_+e_z/\sqrt{2} & e_z^2 & e_ze_-/\sqrt{2}\\ -\frac{1}{2}e_+^2 & e_ze_+/\sqrt{2} & \frac{1}{2}e_+e_- \end{vmatrix}|\langle r\rangle|^2. \quad (8)$$

For our simple case the re-radiation matrix $G_{\mu'\mu}=\sum_{m'}g_{\mu'm'}g_{m'\mu}$ is the same matrix as $F_{\mu\mu'}$ except that all quantities are primed since they refer to the scattered light.

It is convenient for understanding the coherent scattering process to break R up into different $\Delta\mu$ components where $\Delta\mu=|\mu-\mu'|$. Then

$$R(f,g)=R(\Delta\mu=0)+R(\Delta\mu=1)+R(\Delta\mu=2), \quad (9)$$

and

$$R(\Delta\mu=0)=\sum_{\substack{\mu,\mu'\\ \mu=\mu'}}\frac{F_{\mu\mu'}G_{\mu'\mu}}{\Gamma-(i/\hbar)(E_\mu-E_{\mu'})}=\frac{1}{2\Gamma}(1-e_z^2-e_z'^2+3e_z^2e_z'^2), \quad (10)$$

$$R(\Delta\mu=1)=\sum_{\substack{\mu,\mu'\\ |\mu-\mu'|=1}}\frac{F_{\mu\mu'}G_{\mu'\mu}}{\Gamma-(i/\hbar)(E_\mu-E_{\mu'})}=\frac{\frac{1}{2}e_ze_z'}{\Gamma^2+(E_{1,1}-E_{1,0})^2/\hbar}\left[\Gamma(e_-e_+'+e_+e_-')+\frac{i(E_{1,1}-E_{1,0})}{\hbar}(e_-e_+'-e_+e_-')\right]$$
$$+\frac{\frac{1}{2}e_ze_z'}{\Gamma^2+(E_{1,0}-E_{1,1})^2/\hbar}\left[\Gamma(e_-e_+'+e_+e_-')+\frac{i(E_{1,0}-E_{1,-1})}{\hbar}(e_-e_+'-e_+e_-')\right], \quad (11)$$

$$R(\Delta\mu=2)=\sum_{\substack{\mu,\mu'\\ |\mu-\mu'|=2}}\frac{F_{\mu\mu'}G_{\mu'\mu}}{\Gamma-(i/\hbar)(E_\mu-E_{\mu'})}=\frac{1}{4[\Gamma^2+(E_{1,1}-E_{1,-1})^2/\hbar]}$$
$$\times\left[\Gamma(e_-^2e_+'^2+e_+^2e_-'^2)+\frac{i(E_{1,1}-E_{1,-1})}{\hbar}(e_+^2e_-'^2-e_-^2e_+'^2)\right]. \quad (12)$$

For our experimental arrangement $\theta=\theta'=\frac{1}{2}\pi$, $\varphi=0$, $\varphi'=\frac{1}{2}\pi$. Also $E_{1,1}=-E_{1,-1}=g_J\mu_0H$ and $E_{1,0}=0$ so that we may simplify the above equations to

$$R(\Delta\mu=0)=(1/\Gamma)(\tfrac{1}{2}\sin^2\alpha\sin^2\alpha'+\cos^2\alpha\cos^2\alpha'), \quad (13)$$

$$R(\Delta\mu=1)=-\frac{1}{2}\frac{g_J\mu_0H\sin2\alpha\sin2\alpha'}{\Gamma^2+(g_J\mu_0H/\hbar)^2}, \quad (14)$$

$$R(\Delta\mu=2)=-\frac{\Gamma/2\sin^2\alpha\sin^2\alpha'}{\Gamma^2+(2g_J\mu_0H/\hbar)^2}. \quad (15)$$

If the incident light is polarized with $\alpha=\frac{1}{2}\pi$ (i.e., parallel to the Y axis) and the outgoing intensity is

[10] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, 1935), Chap. III.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

measured without polarizers then

$$R(f,g)=\frac{1}{2}\left[\frac{1}{\Gamma}-\frac{\Gamma}{\Gamma^2+(2g_J\mu_0 H/\hbar)^2}\right]. \quad (16)$$

To obtain Eq. (16), the result for the intensity of the scattered light, we have summed Eq. (9) for α' and for $\alpha'+\pi/2$ since this takes account of both orthogonal components of the outgoing beam.

Experimental Details and Results

In Fig. 3 is shown a schematic diagram of the apparatus. The beam source is a stainless-steel oven positioned about 4 in. below the scattering region. The beam effuses through a $\frac{1}{4}$-in.-diam crinkle foil aperture at the top of the oven and is further collimated into a $\frac{1}{2}$-in.-diam region as it passes vertically through the scattering chamber. The beam is then collected on a liquid N_2 trap. A beam flag is provided to determine the degree of instrumental scattering in the absence of the beam and to aid in aligning the optics. Under actual conditions of operation the instrumental scattering can be reduced to several times the dark current of the photomultiplier while the resonance radiation scattered from the beam can easily be made 30 times the dark current. In this arrangement the dc scattered resonance radiation signal is collected by an aluminized light pipe, passed through a narrow-band interference filter for the $^1S_0$–$^1P_1$ resonance line in order to reduce the scattered-light background, and detected by a Dumont 7664 photomultiplier. The photomultiplier output is further amplified by a Hewlett Packard 412A vacuum-tube voltmeter and displayed on a servo recorder.

The externally applied vertical magnetic field at the scattering region was calibrated by proton resonance and found to be 7.51±0.02 G/A. The principal component of the stray magnetic field at the scattering region was 0.35 G in the vertical direction. This field was not compensated but was taken into account in evaluating

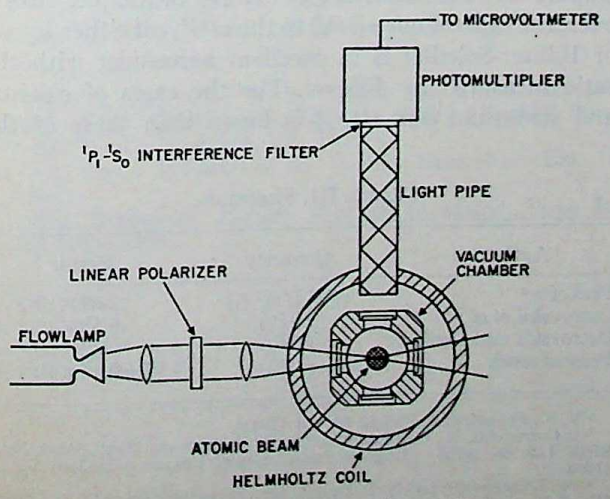


FIG. 3. Schematic diagram of apparatus.

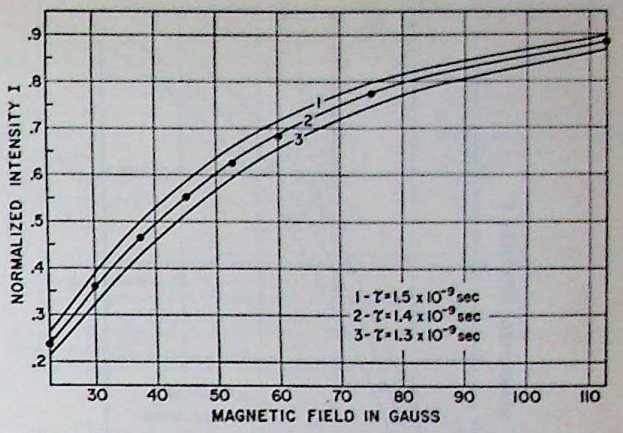


FIG. 4. Normalized plot of level-crossing signal versus magnetic field for zinc. 1, 2 and 3 are theoretical curves. Only the portion of the curve near the half-width is shown.

the data. Most of the measurements were made with the incident light polarized with the electric vector in the direction of observation. Measurements were also made with unpolarized incident light and gave the same results.

The data was taken in the following manner. By means of a reversing switch, the intensity I of the scattered light was recorded successively with the magnetic field in the "up" and then in the "down" direction. A series of about ten such determinations for values of H between 0 and 200 G comprised a run. A plot of I versus H with the zero-field scattered-light intensity subtracted should be an inverted Lorentzian curve. From an approximate value of τ and the value of I at the larger magnetic-field values we may obtain the constant C in Eq. 2. For each run, after C has been determined, the experimental curve is normalized to unity amplitude and plotted together with a family of theoretical Lorentzian curves with different values of τ. In this way τ and an estimate of the uncertainty in τ can be obtained. Between 10 and 15 runs were taken for each element and the values for τ quoted below are the average values derived from these independent runs. Figures 4 and 5 show typical data for zinc and calcium. The slight asymmetry between the field up and the field down points for calcium is due to the incident and observing directions not being exactly perpendicular. Careful arrangement of the optics can eliminate this asymmetry. We find for the lifetime of the $^1P_1$ states the following values: $\tau(\mathrm{Zn})=1.41\pm0.04\times10^{-9}$ sec, $\tau(\mathrm{Ca})=4.48\pm0.15\times10^{-9}$ sec, and $\tau(\mathrm{Sr})=4.97\pm0.15\times10^{-9}$ sec. The quoted errors are twice the standard deviations for the different runs and represent a conservative estimate of the accumulated uncertainty arising from possible instrumental inaccuracies. We have taken $g_J=1.00$ in obtaining τ from the measured linewidths.

The beam density was varied over a factor of about 20 during these measurements in order to detect and correct for coherence narrowing. At the highest beam

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

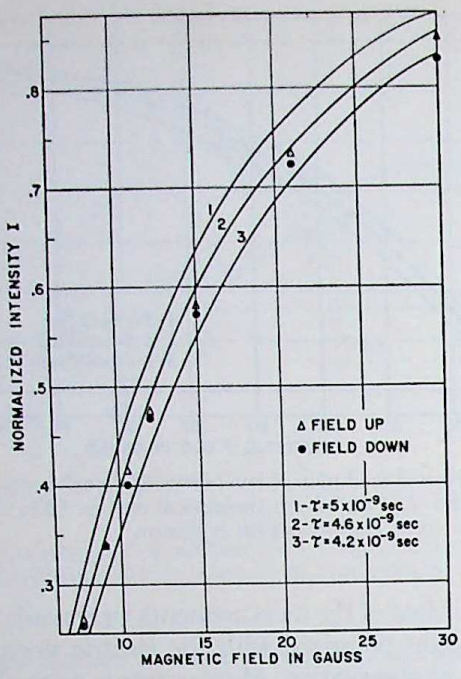


FIG. 5. Normalized plot of level-crossing signal versus magnetic field for calcium. 1, 2, and 3 are theoretical curves.

densities a narrowing of a few percent was obtained. For this reason the results quoted are from those runs taken at low-beam densities.

The possibility exists that self-reversal in the lamp might affect the results. One would expect this effect to distort the Lorentzian line shape. This would occur because as the magnetic field is increased at the scattering atoms, the $\Delta m=\pm 1$ Zeeman levels of the atom would scan a higher intensity part of the lamp profile, and consequently the scattered light would increase faster than it would for a flat profile. No such distortion of the measured Lorentzian shaped lines occurred to within the accuracy of the experiment.

Finally, we wish to mention that originally data for all three elements, with standard deviations of 3–6%, were taken with a somewhat cruder version of the present apparatus. The data presented here were obtained in a physically different location, with entirely different hardware, including different recording instruments and calibrating standards. The results presented here fall within the standard deviations of the preliminary results. This degree of care was taken in part because of disagreement between the preliminary data and previous values in the literature.

TABLE I. Zinc.

Author	Quantity measured	Result
Filippov[a]	$\tau(^3P_.)/\tau(^1P_.)$	6.76×10^{-5}
Auslander[b]	$\tau(^1P_1)$	2.3×10^{-9} sec
Present work	$\tau(^1P_1)$	$1.41\pm0.04\times10^{-9}$ sec
Landman[c]	$\tau(^1P_1)$	$1.38\pm0.05\times10^{-9}$ sec
Billiter-Soleillet[d]	$\tau(^3P_1)$	2.1×10^{-5} sec
Byron[e]	$\tau(^3P_1)$	$2.0\pm0.2\times10^{-5}$ sec
Dumont[f]	$\tau(^3P_1)$	2.9–3.5×10^{-5} sec

[a] A. N. Filippov, Physik Z. Sowjetunion 1, 289 (1932). See also *Optical Transition Probabilities*, a translation of selected Russian articles published for the National Science Foundation, Washington, D. C. by the Israel Program for Scientific Translations, Jerusalem, 1962. In this paper the ratio of the oscillator strengths is given and we have converted to the lifetime ratio.
[b] J. Auslander, Helv. Phys. Acta 11, 562 (1938).
[c] A. Landman and R. Novick, Phys. Rev. 134, A56 (1964).
[d] W. Billeter, Helv. Phys. Acta 7, 841 (1934); P. Soleillet, Compt. Rend. 204, 253 (1937).
[e] F. W. Byron, M. N. McDermott, R. Novick, B. W. Perry, and E. Saloman, Phys. Rev. 134, A47 (1964).
[f] M. Dumont, Thèse de Troisième Cycle, Paris, 1962 (unpublished).

TABLE II. Calcium.

Author	Quantity measured	Result
Prokofev[a]	$\tau(^3P_1)/\tau(^1P_1)$	1.25×10^{-5}
Steinhäuser[b]	$\tau(^1P_1)$	3.5×10^{-9} sec
Ostrovskii *et al.*[c]	$\tau(^1P_1)$	$6.2\pm1\times10^{-9}$ sec
Ostrovskii[d]	$\tau(^1P_1)$	$5.40\pm0.14\times10^{-9}$ sec
Odintsov[e]	$\tau(^1P_1)$	$5.0\pm0.3\times10^{-9}$ sec
Present	$\tau(^1P_1)$	$4.48\pm0.15\times10^{-9}$ sec

[a] V. K. Prokofev, Z. Physik 50, 701 (1928).
[b] A. Steinhäuser, Z. Physik 95, 669 (1935); 99, 300 (1936).
[c] I. Ostrovskii, N. P. Penkin, L. N. Shabanova, Dokl. Akad. Nauk. SSSR 120, 66 (1958) [English transl.: Soviet Phys.—Doklady 3, 538 (1958)].
[d] I. Ostrovskii, N. P. Penkin, Opt. i Spektroskopiya 11, 565 (1961) [English transl.: Opt. Spectry. (USSR) 11, 307 (1961)].
[e] A. I. Odintsov, Opt. i. Spektroskopiya 14, 322 (1963) [English transl.: Opt. Spectry (USSR) 14, 172 (1963)].

Discussion of Results

Tables I, II, and III list the present and previous results for the lifetime of the first excited $^1P_1$ and $^3P_1$ states of zinc, calcium, and strontium. These lifetimes have been calculated from the quoted f values by use of the relation[3]

$$f\tau=\frac{mc}{8\pi^2e^2}\frac{g_2}{g_1}\lambda^2=1.499\frac{g_2}{g_1}\lambda^2.$$

In the case of zinc, one has excellent agreement between the present results and those of Landman although it should be observed that both of these experiments employ the zero-field level-crossing technique. Moreover, the ratio of our $\tau(^1P_1)$ to the $\tau(^3P_1)$ of either Byron or Billiter-Soleillet is in excellent agreement with the ratio measured by Filipov. For the cases of calcium and strontium our $\tau(^1P_1)$ is lower than those of the

TABLE III. Strontium.

Author	Quantity	Result
Prokofev[a]	$\tau(^3P_1)/\tau(^1P_1)$	2.69×10^{-5}
Ostrovskii *et al.*[b]	$\tau(^1P_1)$	6.38×10^{-9} sec
Ostrovskii and Penkin[c]	$\tau(^1P_1)$	6.2×10^{-9} sec
Present work	$\tau(^1P_1)$	$4.97\pm0.15\times10^{-9}$ sec

[a] V. K. Prokofev, Z. Physik 50, 701 (1928).
[b] I. Ostrovskii, N. P. Penkin and L. N. Shabanova, Dokl. Akad. Nauk SSSR 120, 66 (1958) [English transl.: Soviet Physics—Doklady 3, 538 (1958)].
[c] Yu. I. Ostrovskii and N. P. Penkin, Opt. i Spectroskopiya 11, 565 (1961) [English transl.: Opt. Spectry. (USSR) 11, 307 (1961)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IV. Intermediate coupling coefficients.

		Wolfe method	Lifetime ratio method
Zn	c_1	0.5656	0.5657
	c_2	0.8247	0.8246
Sr[a]	c_1	0.5519	0.5526
	c_2	0.8339	0.8334
Ca[a]	c_1	0.5709	0.5717
	c_2	0.8210	0.8204

[a] We use the Prokofev lifetime ratio in the case of Ca and Sr since no direct measurement of $\tau(^3P_1)$ is available for these elements.

Russian measurements, however, we note that their most recent results are in better agreement with ours. We do not know the source of these discrepancies but we should point out the following fact: The recent ingenious improvement in the hook method[11] utilizes total absorption to measure $Nf\Gamma$. The use of this method requires that the absorption follow a Lorentzian law at about three Doppler widths from the line center or typically about 200 times the natural linewidth from the line center. It is important to note that this has never been checked experimentally and any deviation from a Lorentzian line shape well out in the wings of the line could lead to erroneous results.

The excellent agreement in the case of Cd and Zn between the hook measurement of $\tau(^1P_1)/\tau(^3P_1)$ and the recent measurements of $\tau(^1P_1)$ and $\tau(^3P_1)$ for these states indicates that one may get good estimates of $\tau(^3P_1)$ by using our $\tau(^1P_1)$ and the ratio obtained from the hook measurements.

It is also of interest to compare the intermediate coupling coefficients[12] c_1 and c_2 for the $^1P_1$ and $^3P_1$ states as derived from lifetime measurements with those derived from Wolfe's[12,13] theory which makes use of the deviation of the fine structure from an interval rule. This theory takes account of spin-orbit and spin-other-orbit interactions but does not allow for spin-spin interactions or the fact that the $^1P_1$ and $^3P_1$ states can have slightly different radial wave functions. The values of c_1 and c_2 calculated by the two different methods are given in Table IV and are in very good agreement.

[11] Yu. I. Ostrovskii, N. P. Penkin and L. N. Shabanova, Dokl. Akad. Nauk. SSSR **120**, 66 (1958) [English transl.: Soviet Phys.—Doklady **3**, 538 (1958)].

[12] A. Lurio, M. Mandel, R. Novick, Phys. Rev. **126**, 1758 (1962).

[13] H. Kopferman, *Nuclear Moments* (Academic Press Inc., New York, 1958).

TABLE V. Comparison of theoretical and experimental values of the oscillator strengths.

	Ca	Sr	Zn
$\lambda(^1S_0-^1P_1)$	4227 Å	4608 Å	2139 Å
f (exp)[a]	1.80	1.92	1.46
f (B & D)[b]	1.83	1.90	1.49
f (Vainshtein)[c]	2.32		
f (Trefftz)[d]	1.737–1.223		

[a] Present work only.
[b] D. R. Bates and A. Damgaard, Phil. Trans. **242**, 101 (1949).
[c] L. A. Vainshtein, Transactions of the P. N. Lebedev Institute, Vol. XV. (English translation available from Consultants Bureau Enterprises, Inc., New York, 1962.)
[d] E. Trefftz, Z. Astrophys. **29**, 287 (1951).

As pointed out in the introduction, the lifetime of the $^1P_1$ state provides a test of approximate methods used to calculate the oscillator strength of the $^1P_1$ to $^1S_0$ transition. The prediction of the Coulomb approximation method of Bates and Damgaard[6] can be obtained by making use of the formula of Garstang.[14]

$$(^1S_0\|P\|^1P_1)=-\sqrt{2}(c_1+\sqrt{2}c_2)\sigma\cong-\sqrt{6}\sigma,$$

and

$$S(J,J^1)=|-\sqrt{2}(c_1+\sqrt{2}c_2)\sigma|^2\cong 6\sigma^2,$$

where σ^2 is obtained for the p electron from the tables given in the Bates and Damgaard paper. The estimate of the oscillator strengths f from this calculation is shown in Table V. Also included in Table V are other estimates for f using explicit forms for the wave functions of the two states. The surprising conclusion to be drawn from Table V is that the Coulomb approximation is in very good agreement with the experimental results.

Note added in proof. After this article was submitted for publication a paper with the following reference appeared: E. Hulpke, E. Paul, and W. Paul, Z. Physik **177**, 257 (1964). In this paper the following results are given: $\tau(\text{Ca})=4.67\pm0.11\times10^{-9}$ sec and $\tau(\text{Sr})=4.56\pm0.21\times10^{-9}$ sec, which are in excellent agreement with ours. As in their paper we have also considered the transitions from the $^1P_1$ to the $^1D_2$ state to be negligible compared to the decay to the ground state.

ACKNOWLEDGMENTS

We would like to thank Robert Holohan for his help in assembling the apparatus and in taking much of the final data. We also appreciate several helpful discussions with Professor R. Novick and A. Gallagher.

[14] R. H. Garstang, J. Opt. Soc. Am. **52**, 845 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

g_J of the $(5s5p)\,^3P_1$ Level of Cd and the $(6s6p)\,^3P_1$ Level of Hg by High-Field Double Resonance*

R. KOHLER

Columbia Radiation Laboratory, Columbia University, New York, New York and Physics Department, New York University, University Heights, New York, New York,

AND

P. THADDEUS

Goddard Institute for Space Studies, National Aeronautics and Space Administration, New York, New York

(Received 20 January 1964)

An optical double resonance experiment at high magnetic fields has determined g_J for the $(5s5p)\,^3P_1$ level of the even cadmium isotopes to be 1.499846(13), and g_J for the $(6s6p)\,^3P_1$ level of the even mercury isotopes to be 1.486094(8). Resonance was observed in all cases at a frequency of exactly 24 Gc/sec, and magnetic fields near 11 430 G. At these high fields the Zeeman energy is of the order of 10^{-3} of the fine structure separations of the triplet terms of the $(nsnp)$ configuration, and the $\Delta m = \pm 1$ transitions are split by 9.51(7) G for Cd, and by 2.99(7) G for Hg. This splitting represents several linewidths in the case of Hg, and many linewidths in that of Cd. The average field of the two $\Delta m = \pm 1$ transitions, however, determines g_J to high precision independent of second-order corrections.

I. INTRODUCTION

THE double resonance[1,2] and level crossing techniques[3,4] have been used in extensive studies of hyperfine structure (hfs) in the lowest lying $(nsnp)\,^3P_1$ level of mercury,[5–11] cadmium,[12–17] and zinc.[18–20] Sufficient sensitivity has been obtained to allow the determination of the hfs, and consequently the nuclear moments, of radioactive isotopes with half-lives as short as a few hours.[15,16]

While the level crossing method has the great advantage in practice of doing away with all radio-frequency and microwave apparatus, aside from nuclear magnetic-resonance (NMR) magnetometers, it allows a determination of only the ratio of the hyperfine interaction constants to g_J, and an independent measurement of g_J must therefore be made to determine these constants to high precision.[16] While g_J can of course be measured by double resonance in the odd isotopes at various fields, or by observing both double resonance and level crossings,[9,13] it is most directly found from double resonance in the even isotopes of the element.

In principle, the higher the frequency and the greater the magnetic field, the more accurate the determination of g_J. In a series of experiments designed to measure g_J of the lowest lying $^3P_1$ level of zinc, cadmium, and

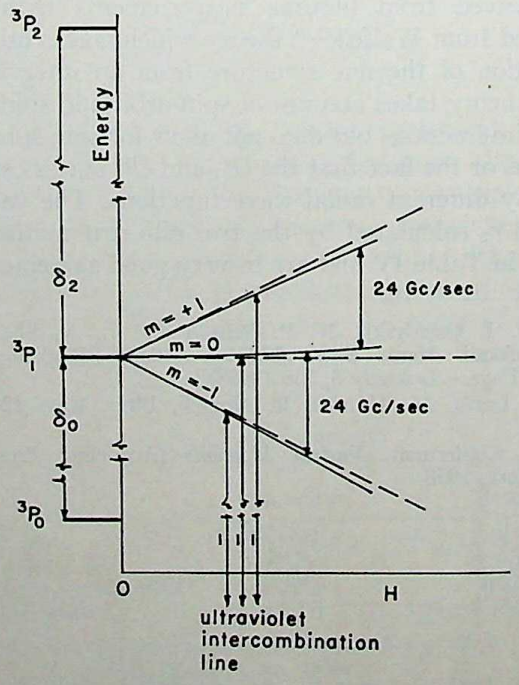


FIG. 1. Zeeman effect of the $^3P_1$ state of the even isotopes of zinc, cadmium and mercury, showing the effect of second-order interaction with the other fine structure levels of the configuration.

* Work supported by the Joint Services (the U. S. Army, the U. S. Navy Office of Naval Research, and the U. S. Air Force Office of Scientific Research) and by the U. S. Office of Naval Research under contract Nonr-266(45).

[1] J. Brossel and F. Bitter, Phys. Rev. 86, 308 (1952).
[2] G. W. Series, Rept. Progr. Phys. 22, 280 (1959).
[3] F. D. Colegrove, P. A. Franken, R. R. Lewis, and R. H. Sands, Phys. Rev. Letters 3, 420 (1959).
[4] P. A. Franken, Phys. Rev. 121, 508 (1961).
[5] J. N. Dodd, Proc. Phys. Soc. (London) 77, 669 (1961).
[6] R. H. Kohler, Phys. Rev. 121, 1104 (1961).
[7] H. R. Hirsch, J. Opt. Soc. Am. 51, 1192 (1961).
[8] C. Brot, J. Phys. Radium 22, 412 (1961).
[9] R. D. Kaul, Ph.D. thesis, Case Institute of Technology, 1963 (unpublished).
[10] W. Smith, Ph.D. thesis, Massachusetts Institute of Technology, 1963 (unpublished).
[11] C. V. Stager, Phys. Rev. 132, 275 (1963).
[12] R. F. Lacey, Ph.D. thesis, Massachusetts Institute of Technology, 1959 (unpublished).
[13] P. Thaddeus and R. Novick, Phys. Rev. 126, 1774 (1962).
[14] M. N. McDermott and R. Novick, Phys. Rev. 131, 707 (1963).
[15] F. W. Byron, M. N. McDermott, and R. Novick, Phys. Rev. 132, 1181 (1963).
[16] P. Thaddeus and M. N. McDermott, Phys. Rev. 132, 1186 (1963).
[17] M. N. McDermott, R. Novick, B. W. Perry, and E. Saloman, Phys. Rev. 134, B25 (1964).
[18] V. K. Böckman, H. Krüger, and E. Rechnagel, Ann. Physik 20, 250 (1957).
[19] F. W. Byron, M. N. McDermott, R. Novick, B. W. Perry, and E. Saloman, Phys. Rev. 134, A47 (1964).
[20] A. Landman and R. Novick, Phys. Rev. 134, A56 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Atomic constants of Zn, Cd, and Hg. $\lambda(^3P_1)$ and $\lambda(^1P_1)$, the wavelengths of the resonance lines of the lowest lying $(nsnp)$ configuration, and δ_2 and δ_0, the fine structure separations of the triplet terms, are taken from C. E. Moore.[a]

Atom	τ $(^3P_1)$ (sec)	τ $(^1P_1)$ (sec)	$\lambda(^3P_1)$ (Å)	$\lambda(^1P_1)$ (Å)	δ_2 (cm^{-1})	δ_0 (cm^{-1})
Zn	$2.00(20)\times10^{-5}$ [b]	$1.38(5)\times10^{-9}$ [c]	3077	2139	388.927	190.082
Cd	$2.39(4)\times10^{-6}$ [d]	$1.66(5)\times10^{-9}$ [e]	3261	2288	1170.866	542.113
Hg	1.18×10^{-7} [f]	0.3–1.6×10^{-9} [g]	2537	1849	4630.677	1767.220

[a] C. E. Moore, *Atomic Energy Levels* (U. S. Government Printing Office, 1949–58).
[b] See Ref. 19.
[c] See Ref. 20.
[d] See Ref. 23.
[e] A. Lurio and R. Novick, Phys. Rev. 134, A608 (1964).
[f] See the last paper of Ref. 22.
[g] A. C. G. Mitchell and M. W. Zemansky, *Resonance Radiation and Excited Atoms* (Cambridge University Press, Cambridge, 1961), 2nd ed., p. 147.

mercury to high precision (it is $\frac{3}{2}$ on the basis of LS coupling), we therefore chose to work at a frequency of 24 Gc/sec, and a field near 11 400 G—approaching the limit attainable with the available 12-in. electromagnet when operating with a gap of 2 in.

The $\Delta m=\pm1$ transitions for these states coincide at low magnetic fields, but at high fields interaction with the neighboring fine structure terms of the configuration (and to a negligible degree with levels of other configurations) splits the resonances into two well-resolved components, as shown in Fig. 1. Precise measurement of the splitting of the two resonances therefore serves also as a check on the theory of the second-order fine structure interactions.

II. THEORY

Signal Strengths

It is well known that double resonance can be observed between hyperfine or Zeeman states excited by resonance radiation in a variety of ways. Population differences can be produced by resonance radiation that is polarized, nonisotropic, or confined to a narrow wavelength interval, and all of these methods have been put to practical use.[1,2,6]

To dispense with ultraviolet polarizers, we relied on the latter two effects and used the experimental geometry shown in Fig. 2. Only lamps containing the natural metal were used. In the case of cadmium, the Zeeman shift of the $m=\pm1$ states at 11 400 G is somewhat greater than the isotope shifts and the hyperfine splittings of the natural cadmium isotopes (and the optical Doppler width),[14] and only the $m=0$ state is therefore illuminated. It is then readily calculated in terms of the transition probabilities that at double resonance the resonance fluorescence at 90° is reduced by a factor of $\frac{1}{4}$.

For mercury, on the other hand, the isotope shifts and hyperfine splittings are comparable to the Zeeman shift of the $\Delta m=\pm1$ states,[21] and we must rely mainly on the nonisotropic nature of the resonance radiation. In the least favorable case, on the assumption that all three states are equally illuminated, it is calculated that the resonance fluorescence at 90° is reduced by a factor of $\frac{1}{12}$ at double resonance.

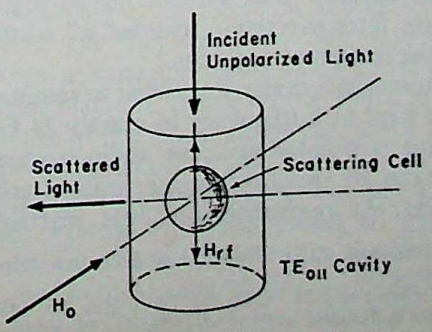


FIG. 2. Geometry of the experiment.

Linewidths

Several effects determine the linewidth and shape of the double resonance effect, and have all been considered in detail elsewhere. Most fundamentally, the natural lifetime τ of the state gives a Lorentzian line shape with full width at half-maximum of $g_J\mu_0\Delta H_n=\Delta\nu_n=1/\pi\tau$.[1,6] The natural linewidths for zinc, cadmium, and mercury calculated from the lifetimes in Table I are listed in Table II.

If the rf magnetic field is strong enough to induce many transitions during the lifetime τ, the resonance will be power broadened.[1,6] In practice the rf power is always attenuated to the point where this effect is small, and it will therefore not be considered further.

Double resonance owes its high resolution to the fact that Doppler broadening is determined by the radio rather than by the optical frequency. At microwave frequencies, however, the Doppler linewidth $\Delta\nu_d=(\nu/c)\times(8NkT\ln2/M)^{1/2}$, where N is Avogadro's number and

TABLE II. Theoretical and observed linewidths in milligauss. ΔH_{calc} is taken as the sum of the natural, Doppler, magnetic field, and wall collision widths.

Atom	ΔH_n	ΔH_d	ΔH_m	ΔH_w	ΔH_{calc}	ΔH_{obs}
Zn	8	29	50	37	124	...
Cd	68	18	50	26	162	175
Hg	1295	10	50	15	1370	1300

[21] F. Bitter, S. P. Davis, B. Richter, and J. E. R. Young, Phys. Rev. 96, 1531 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

M is the molecular weight, may not be altogether negligible, as is shown in Table II.

Inhomogeneity of the magnetic field over the scattering sample contributes a linewidth ΔH_m that can often be reduced by careful alignment of the pole pieces. We found, however, that as our magnetic field approached 11 400 G, its homogeneity fell off rapidly due presumably to uneven saturation. We estimate that $\Delta H_m = 50$ mG over a 5-mm sample.

Collisions with the walls of the scattering cell effectively shorten the lifetime of the excited state, and contribute a wall-quenching linewidth $\Delta\nu_w$ which depends mainly on the size of the cell. We have made estimates of $\Delta\nu_w$ in Table II on the assumption that $\Delta\nu_w = 1/\pi\tau_w = (3NkT/M)^{1/2}/\pi a$, where τ_w is the time for an atom at the temperature of the optimum vapor pressure to travel the radius a of the scattering cell.

Collisions between the scattering atoms will broaden the resonance at high vapor densities, while multiple scattering of the optical photons may actually decrease the linewidth (coherence narrowing).[22,23] Both of these effects may be avoided by working at low enough temperatures.

Quenching by atomic or molecular contaminants in the scattering cell, on the other hand, may contribute a linewidth $\Delta\nu_c$ which cannot be easily estimated in advance, and cannot be reduced once the cell is made. We shall see that possibly except for zinc we found no evidence of quenching of this kind.

Zeeman Effect

Interaction with the two neighboring triplet levels of the ($nsnp$) configuration contributes a term which is of the order of the (Zeeman energy)^2/(fine structure energy) $\approx 10^{-3} \times$ (Zeeman energy) to the energy of the $m = \pm 1$, 0 states. More precisely, from perturbation theory,[24]

$$W_{+1} = g_J\mu_0 H - \frac{\alpha^2(g_S - g_L)^2\mu_0^2 H^2}{4\delta_2}, \tag{1a}$$

$$W_0 = \alpha^2(g_S - g_L)^2\mu_0^2 H^2\left(\frac{2}{3\delta_0} - \frac{1}{3\delta_2}\right), \tag{1b}$$

$$W_{-1} = -g_J\mu_0 H - \frac{\alpha^2(g_S - g_L)^2\mu_0^2 H^2}{4\delta_2}, \tag{1c}$$

where μ_0 is the Bohr magneton, α is the mixing coefficient which indicates the departure from Russell-Saunders coupling, and δ_0 and δ_2 are the magnitude of the fine structure separations. We adopt the convention that g_J is positive. Interaction with the singlet term of the configuration, and with terms of other configurations, contributes less than one part in 10^6 to the energy, and third-order terms vanish.

For double resonance at a fixed frequency ν_{dr}, the field H^+, for which $m = 0 \rightarrow +1$, is greater than H^-, for which $m = 0 \rightarrow -1$. To within one part in 10^6 we find from Eq. (1) that in terms of the average field $\bar{H} = (H^+ + H^-)/2$

$$g_J = h\nu_{dr}/\mu_0\bar{H}. \tag{2a}$$

This becomes

$$g_J = g_p(\nu_{dr}/\bar{\nu}_p), \tag{2b}$$

if the field is measured in terms of the NMR frequency ν_p of protons. g_p is defined such that the proton magnetic dipole-moment operator is $\boldsymbol{\mu}_p = g_p\mu_0\mathbf{I}_p$. For protons in mineral oil we will take $g_p = 3.041978(3) \times 10^{-3}$, from the average of the measurements of g_s/g_p for protons in mineral oil of Koenig, Prodell, and Kusch,[25] and Beringer and Heald,[26] and the recent measurement of g_s of Wilkinson and Crane.[27]

For the separation of the two resonances we find that

$$\Delta H = H^+ - H^- = \frac{\alpha^2(g_S - g_L)^2\mu_0\bar{H}^2}{g_J}\left(\frac{4}{3\delta_0} - \frac{1}{6\delta_2}\right), \tag{3a}$$

where δ_0 and δ_2 are in ergs. When the field is measured in terms of the proton frequency,

$$\Delta\nu_p = \frac{\alpha^2(g_S - g_L)^2\bar{\nu}_p^2}{cg_pg_J}\left(\frac{4}{3\delta_0} - \frac{1}{6\delta_2}\right), \tag{3b}$$

where δ_0 and δ_2 are in cm^{-1}.

III. EXPERIMENT

A schematic illustration of the experiment is shown in Fig. 3. Resonance radiation from an rf lamp was focused with a quartz condensing lens onto a small cell containing cadmium or mercury vapor. The cell was located within a microwave cavity between the poles of a 12-in. electromagnet. This cavity in the case of cadmium also served as an oven to control the vapor density in the scattering cell.

Resonance fluorescence of the atomic vapor was observed with a photomultiplier at an angle of 90°. A light pipe allowed the photomultiplier to be removed from the magnetic field without loss of solid angle. A glass color filter effectively removed all scattered light except the resonance radiation.

The microwave cavity was tuned to resonance with a klystron kept at a frequency of exactly 24 Gc/sec, the field being modulated at a low audio frequency, typically 37 cps, by small rectangular coils attached to the pole faces. At resonance, the light scattered at 90°

[22] J. P. Barrat, J. Phys. Radium **20**, 541, 633, 657 (1959).
[23] F. W. Byron, M. N. McDermott, and R. Novick, Phys. Rev. **134**, A615 (1964).
[24] A. Lurio, M. Mandel, and R. Novick, Phys. Rev. **126**, 1758 (1962).
[25] S. H. Koenig, A. G. Prodell, and P. Kusch, Phys. Rev. **88**, 191 (1952).
[26] R. Beringer and M. A. Heald, Phys. Rev. **95**, 1474 (1954).
[27] D. T. Wilkinson and H. R. Crane, Phys. Rev. **130**, 852 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

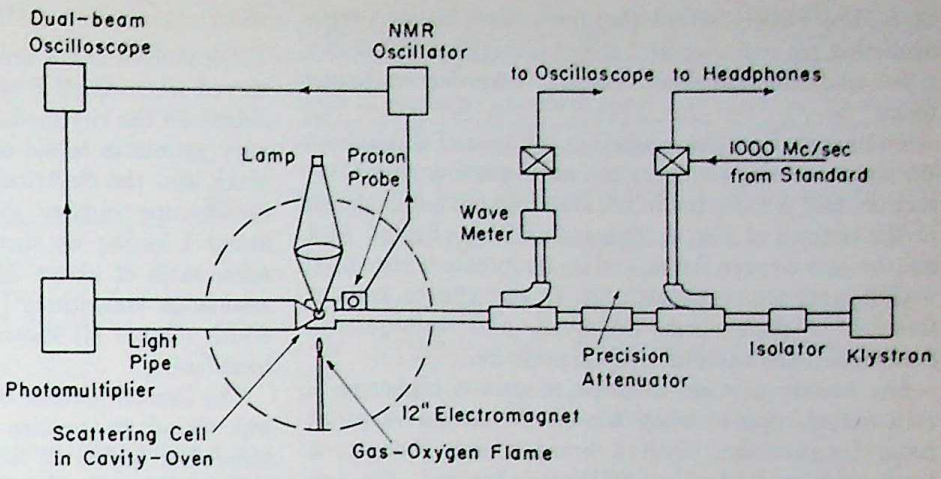


FIG. 3. Schematic diagram of the experiment.

changed in intensity, and the signal from the photomultiplier was displayed on an oscilloscope.

Field measurements were made with an NMR magnetometer, whose small mineral-oil probe was located as close to the scattering cell as possible. The offset correction was measured with a second probe and oscillator.

Optics

The cadmium lamp used consisted of a small quartz cylinder, 1.5 cm in diameter and 2.5 cm long, carefully baked out, purged with an rf discharge in argon and then filled with a small amount of distilled natural cadmium and about 1 torr of spectroscopically pure argon. It was identical to the lamps previously used in this laboratory for several level crossing and double resonance experiments.[13,14] It was excited by being placed, without cooling, in the tank coil of a Hartley oscillator operating in the vicinity of 30 Mc/sec. Experience has shown that these lamps furnish about 1 W (total output over 4π solid angle) in the 3261-Å intercombination line.

For mercury an Osram lamp was used that had the outer jacket removed, and was cooled with a gentle air stream to eliminate self-reversal. It too was excited with a radiofrequency oscillator operating near 30 Mc/sec and was run at the same current specified by the manufacturer for operation at 60 cps.

The scattering cells were designed to have as small an effect as possible on the electrical properties of the microwave cavity. They were very thin walled quartz bubbles, 5 mm in diameter, blown at the end of a length of 2 mm outer diameter quartz tubing, which, when cut off at a length of 5 cm, served as the tail of the cell. These cells were baked out at 1000°C for a day or longer, purged with a Tessla coil for several minutes, and filled with a small quantity of distilled metal, all under very high vacuum (less than 10^{-8} Torr for the last stage). The light pipe used to remove the RCA-1P28 photomultiplier from the vicinity of the magnet was a glass tube aluminized on the inside, 1 in. in diameter, and about $2\frac{1}{2}$ ft long. To pass the 3261-Å line of cadmium, or the 2537-Å line of mercury, and exclude other scattered light, a Schott UG-11 or UG-5 colored glass filter was cemented to the end of the light pipe.

Microwave Cavity and Oven

The TE_{011} cylindrical aluminum cavity oven used for cadmium is shown in Fig. 4. It had an internal diameter of 1.4 cm, a Q of about 4000 when loaded with the cell, and was coupled about 50% to the microwave line. A hole 6 mm in diameter at the top of the cavity, below cutoff at 24 Gc/sec, allowed illumination of the scattering cell. Small slits in the side of the cavity, also below cutoff, allowed the light scattered at 90° to be observed.

The waveguide connecting the cavity to the microwave line was thinned to provide thermal insulation. The cavity was heated with a small gas-oxygen flame, the temperature being monitored by a small copper-constantin thermocouple. Thermal regulation was provided

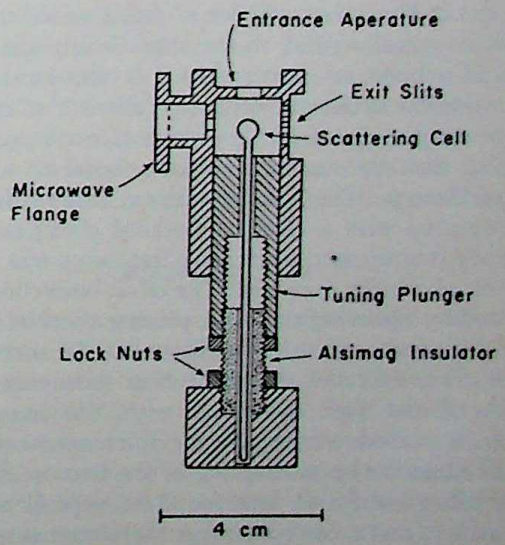


FIG. 4. The cavity oven used for cadmium.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

by a Honeywell R7086A thermocouple-activated regulator that played a stream of cool air against the cavity when the thermocouple voltage exceeded a preset figure.

The long tail of the scattering cell passed through an aluminum tuning plunger, and an insulating refractory section, and terminated in the aluminum cylinder shown in the bottom of Fig. 4. This cylinder was heated with another gas-oxygen flame, and its temperature regulated with a separate regulator unit. It was always kept at least 40°C cooler than the oven, and consequently determined the cadmium vapor pressure.

For mercury, in an attempt to gain a higher Q, a thin-walled copper cavity was used that was kept at room temperature. With a brass tuning plunger, a loaded cavity Q of about 8000 was obtained. The tail of the cell was usually kept at room temperature, but could also be cooled in an ice bath.

Microwave Measurements

The Varian VA-98E klystron used in this experiment is an extremely stable tube that oscillates only over the region from 23.6 to 24.4 Gc/sec, with an average output of about 30 mW. To achieve maximum stability, we used batteries for both the heater and repeller voltages, and a well-regulated power supply for the beam voltage. No electronic stabilization of the tube was employed. During a run it was kept manually to within an audio-beat note of the 24th harmonic of the 1000 Mc/sec output of the laboratory's frequency standard. This represents a stability of about one part in 10^6, and is therefore quite adequate for the purpose of this experiment.

Field Measurements

The proton probe of the NMR magnetometer used for field measurements consisted of a small spherical cavity filled with mineral oil drilled in the center of a teflon spool. The inductance for a transistorized marginal oscillator was wound on the spool which was then placed in a brass or copper shield. It was located as close as possible to the scattering cell (about 4 cm away), and the output signal of the marginal oscillator was displayed with the double resonance signal on a dual beam oscilloscope. The NMR frequency, near 48 Mc/sec, was measured with a Hewlett-Packard direct reading frequency counter whose reference frequency was taken from the laboratory standard. The offset correction was measured by removing the oven, placing a second probe driven by another identical oscillator at the scattering cell location, and measuring the beat frequency ($\approx$4 kc/sec) of the two oscillators with the frequency counter. A pentode mixer gave the difference frequency without excessive cross coupling of the two oscillators.

The offset correction was found to depend on the field history, and a correction was therefore made after each field measurement.

IV. RESULTS

For cadmium the onset of power broadening was observed when about 5 mW of microwave power were incident on the cavity, in rough accord with our preliminary estimates based on the lifetime of the $(5s5p)\ ^3P_1$ level, and the electrical properties of the cavity. With oscilloscope display giving a detector bandwidth of about 1 kc/sec we succeeded in obtaining a signal-to-noise ratio of about 30 for both the $\Delta m=\pm 1$ double resonance transitions [Fig. 5(a)]. The observed linewidth (Table II) shows no evidence of contamination broadening.

The shorter lifetime of the $(6s6p)\ ^3P_1$ level of mercury was found to require for maximum signal strength somewhat higher rf fields than the VA-98E klystron and aluminum cavity could provide. A higher Q copper cavity subsequently used for mercury just allowed the level of power broadening to be reached. The observed resonances [Fig. 5(b)] had a signal-to-noise ratio of about 5. It should be recalled that because of the illumination of the $m=\pm 1$ states the mercury signals are expected to be about 3 times weaker than those of cadmium. The additional discrepancy in signal strength can be attributed to the somewhat weaker output of the mercury lamp. The linewidth was in agreement with expectations (Table II).

We had originally hoped to extend our measurements to the $(4s4p)\ ^3P_1$ level of the even zinc isotopes, although we were aware that the long lifetime of this level would produce serious quenching of the excited state atoms against the wall of the scattering cell. We, in fact, made several runs using narrow-band phase-sensitive detection with time constants as long as 5 sec, but failed to detect either double resonance or fluorescence. Beyond the quenching and broadening mechanisms listed in Table II, we are inclined to attribute this to contaminants produced by the high cell temperature required for zinc.

The magnetic fields at which double resonance was observed for cadmium and mercury are listed in Table III. From Eq. (2) we then find for the $(5s5p)\ ^3P_1$ level of cadmium that

$$g_J=1.499846(13)=\tfrac{3}{2}-154(13)\times 10^{-6},$$

in good agreement with the values $g_J^{111}=\frac{3}{2}-171(9)\times 10^{-6}$ and $g_J^{113}=\frac{3}{2}-157(9)\times 10^{-6}$ found for the odd isotopes from level crossings[13] and zero-field double resonance[12] (we have here recalculated g_J^{111} and g_J^{113} taking $g_p=3.041978\times 10^{-3}$).

Likewise, we find for the $(6s6p)\ ^3P_1$ level of mercury that

$$g_J=1.486094(8)=\tfrac{3}{2}-13\,906(8)\times 10^{-6}.$$

This is in reasonably good agreement with a recent preliminary value of $g_J^{199}=1.486118(16)$ found by Smith,[28] who considered both his own level-crossing

[28] W. Smith (private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

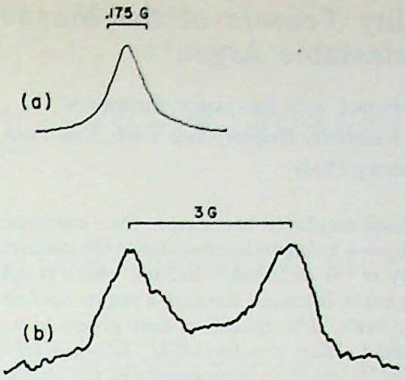


FIG. 5. Oscilloscope traces of the double resonance signals: (a) is one of the $\Delta m = \pm 1$ cadmium transitions, (b) is the sum of these two transitions for mercury, split by the second-order interactions.

study of Hg^{199}, and the results of low-field double resonance.[11] It is also in agreement with the less precise value 1.486350(300) found by Dodd[29] from double resonance in the even isotopes. It is in disagreement, however, with the self-consistent values $g_J^{199} = 1.486147(10)$ and $g_J^{201} = 1.486156(18)$ calculated by Kaul in his thesis[9] from his study of level crossings, and the low-field double resonance results.[6,11]

For cadmium the known lifetimes of the $^3P_1$ and $^1P_1$ levels allow α to be calculated directly from the relation

$$\frac{\beta^2}{\alpha^2} = \frac{\tau(^1P_1)\lambda^3(^3P_1)}{\tau(^3P_1)\lambda^3(^1P_1)}, \tag{4}$$

and the normalization condition $\alpha^2+\beta^2=1$. Using the lifetimes and wavelengths of Table I we find that $\alpha = 0.99900(5)$ and $\beta = -0.0448(11)$. The phase convention of Condon and Shortley is used to determine the sign of β.[24] From Eq. (3) the splitting of the resonances is then calculated to be $\Delta\nu_p = 40.245(4)$ kc/sec, or $\Delta H = 9.452(1)$ G. This is seen from Table III to be 0.255 kc/sec lower than the experimental value, but just within the experimental uncertainty.

Knowing α and β independently from the lifetimes, we are in a position to determine the sum of the relativistic and diamagnetic corrections to g_J. Theoretically

$$g_J = \tfrac{1}{2}\alpha^2(g_L+g_S)+\beta^2 g_L+\Delta g_{\rm rel}+\Delta g_{\rm diam}. \tag{5}$$

Taking $g_L = 1$,[30] and $g_S = 2.002319$,[27] we find that[31]

$$\Delta g_{\rm rel}+\Delta g_{\rm diam} = -318(40)\times 10^{-6}.$$

Since for mercury the lifetime of the $^1P_1$ level is poorly known, the sum of the relativistic and diamagnetic corrections cannot be found. Instead, however, we may use Eq. (5) neglecting these corrections to obtain $\alpha = 0.9849(4)$ and $\beta = -0.1733(24)$. We then calculate from Eq. (3) the splitting of the mercury resonances to be $\Delta\nu_p = 12.468(8)$ kc/sec or $\Delta H = 2.928(2)$ G. It is interesting to note that this is lower than the experimental value by 0.252 kc/sec, nearly the same as in the case of cadmium. Since the measurement of the splitting ΔH is not susceptible to the same systematic errors as the absolute field measurements, the uncertainties in ΔH given in Table III must be considered quite conservative, and this discrepancy may be significant.

TABLE III. Measured fields for double resonance in cadmium and mercury, in terms of the NMR frequency of protons in mineral oil, and in G, taking for $1/2\pi\times$ the gyromagnetic ratio 4.25760(4) kc/sec-G.[a] A series of runs were made for each atom, care being taken to randomize systematic errors from run to run by slightly relocating the cell and probe within the field, reversing the polarity of the magnet, varying the modulation frequency, changing the vapor pressure, reassembling the optics, etc. The uncertainties quoted are 3 times the standard deviation of the means of the individual runs.

Atom	H^+	H^-	ΔH
Cd	48 696.91(66) kc/sec	48 656.40(45) kc/sec	40.50(28) kc/sec
	11 437.64(19) G	11 428.13(15) G	9.51(7) G
Hg	49 133.45(35) kc/sec	49 120.73(38) kc/sec	12.72(26) kc/sec
	11 540.17(14) G	11 537.19(14) G	2.99(7) G

[a] R. L. Driscoll and P. L. Bender, Phys. Rev. Letters 1, 413 (1958).

V. DISCUSSION

While an accuracy of about one part in 10^5 has been attained in the field measurements reported in this experiment, for cadmium this represents an uncertainty of one linewidth, and can certainly be improved upon. In retrospect we believe that higher precision could have been realized at somewhat lower frequencies and fields. Most of our uncertainty can be ascribed to the inhomogeneities which crept in as the field approached 11 400 G. Had we worked at half the present field and frequency, for example, the field would have been considerably more uniform, the scattering cell would have been twice as large (and easier to construct and more durable), the Doppler and wall collision linewidths halved, and the exit and entrance apertures increased in area by four. It is probable that under these circumstances we would have had sufficient signal strength to observe double resonance in zinc, and that the cadmium resonance could have been measured to a small fraction of its width.

ACKNOWLEDGMENTS

We wish to express our thanks to the staff of the Columbia Radiation Laboratory for their invaluable assistance, and to Professor R. Novick for his advice and support. We would like especially to acknowledge the assistance of H. Feldman, who collaborated with us during a part of the cadmium experiment. Finally, Raymond Eisenstark assisted us liberally during all phases of this work as data taker and calculator, and T. Psaropulos kindly prepared the drawings.

[29] J. N. Dodd, Proc. Phys. Soc. (London) 78, 65 (1961).

[30] This neglects effects due to nuclear motion, which are expected to contribute a correction to g_L of the order of a few ppm. See A. Abragam and J. H. Van Vleck, Phys. Rev. 92, 1448 (1953).

[31] The theoretical values of $g_J = \frac{3}{2} - 18(80)\times 10^{-6}$ and $\Delta g_{\rm rel} + \Delta g_{\rm diam} = -135(81)\times 10^{-6}$ given in Ref. 13 are in error. These values were calculated from Eq. (13), Ref. 13 on the basis of outdated lifetimes—not those given in Table I, Ref. 13. To the best of our knowledge, the remaining results in that work are self-consistent.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Determination of the Polarizability Tensors of the Magnetic Substates of $^3P_2$ Metastable Argon*†

EDWARD POLLACK,‡ EDWARD J. ROBINSON,§ AND BENJAMIN BEDERSON
Department of Physics, New York University, University Heights, New York, New York

(Received 8 January 1964)

The atomic beam E-H gradient-balance method has been employed to measure the z-component $\alpha_z(m_J)$ of the electric polarizability tensors of the $m_J=+2$ and $m_J=+1$ magnetic substates of $^3P_2$ metastable argon. The measurements were normalized to the polarizability of $^3S_1$ metastable helium, which is calculable to within an accuracy of 1%. In this method, a metastable beam is passed through a region containing transverse, congruent, inhomogeneous electric and magnetic fields. The condition that atoms in a particular magnetic substate experience no deflection in such a field region is $\alpha_z(m_J)E(\delta E/\delta z)=\mu_{\rm eff}(\delta H/\delta z)$, where $\mu_{\rm eff}$ is the magnetic moment in the field H, and $\delta E/\delta z$, $\delta H/\delta z$ are the z components of the gradients of the magnitudes of the electric and magnetic fields, respectively. Because of the congruency of the E and H fields $(\delta E/\delta z)/E=(\delta H/\delta z)/H$, and therefore $\alpha_z(m_J)=\mu_{\rm eff}H/E^2$. Because of the normalization, it was unnecessary to find either E or H. The results, in units of 10^{-24} cm^3, are $\alpha_z(\pm1)=50.4\pm3.5$ and $\alpha_z(\pm2)=44.5\pm3.1$. The measured values of $\alpha_z(+1)$ and $\alpha_z(+2)$ can be used to determine the polarizability tensors of each of the magnetic substates. The other components of the diagonalized tensors, with $\alpha_x(m_J)=\alpha_y(m_J)$, in units of 10^{-24}cm^3 are $\alpha_x(0)=46.4\pm2.4$, $\alpha_x(\pm1)=47.2\pm2.3$, $\alpha_x(\pm2)=50.4\pm3.5$, and $\alpha_z(0)=52.4\pm4.8$.

INTRODUCTION

RECENTLY an atomic beam method has been developed to measure the electric polarizabilities of atoms.[1–3] Using this technique, called the "E-H gradient balance method," polarizability determinations are independent of the atomic-beam velocity distribution and specific values of the field gradients and apparatus geometry except where employed in the calculation of the electric field at the beam position from knowledge of the applied voltage and pole-face contour. Salop, Pollack, and Bederson[2,3] used this method to measure the polarizabilities of the alkalis, and the results of this work, as well as those of Chamberlain and Zorn,[4] are in agreement with recent theoretical determinations, to within experimental error.[5,6] For all these systems, the ground states are S states and the polarizability is therefore scalar.

We will here briefly review our basic technique, applied to an atom possessing a tensor polarizability. Consider an atom traversing a region of space containing both an inhomogeneous transverse electric field and an inhomogeneous transverse magnetic field, with both fields possessing large transverse gradients. We assume that these fields are congruent, that is, the electric and magnetic fields as well as their gradients are proportional and in the same direction throughout the interaction region. The atom experiences a transverse force due to its magnetic moment given by

$$F_m=\mu_{\rm eff}(m_J)(\delta H/\delta z)\,, \tag{1}$$

where $\mu_{\rm eff}(m_J)$ is the effective atomic magnetic moment of the magnetic substate m_J and $\delta H/\delta z$ is the transverse component of the gradient of the magnitude of the magnetic field.

For an atomic system possessing a tensor polarizability the electric dipole moment $\boldsymbol{\mu}_e$ induced in the atom by the electric field $\mathbf{E}$ is generally not along the field direction. For the case where the polarizability tensor is in diagonal form, $\boldsymbol{\mu}_e$ is given by

$$\boldsymbol{\mu}_e=\hat{\imath}\alpha_xE_x+\hat{\jmath}\alpha_yE_y+\hat{k}\alpha_zE_z\,,$$

where α_x is the xx component of the polarizability tensor α, etc. For S-state atoms, $\alpha_x=\alpha_y=\alpha_z=\alpha$, i.e., the polarizability is a scalar. For the magnetic substates of P-state atoms, this is in general not the case and the polarizability of each of the substates is a tensor. In the present work, the electric and magnetic fields, which are in the z direction, establish the z axis as a symmetry axis in the atom. Thus, α is in diagonal form, with the x and y components being equal. In terms of matrix elements, the polarizability components may be

* This work was supported in part by the Advanced Research Projects Agency, through the U. S. Office of Naval Research, Contract N-ONR-285(15) and by the Defense Atomic Support Agency, through the U. S. Army Research Office, Grant DA-ARO-(D)-31-124-G269, Durham, North Carolina.

† From part of a thesis submitted by E. Pollack to the Graduate Faculty of New York University in partial fulfillment of the requirements for the Ph.D. degree, October 1963.

‡ Present address: Physics Department, University of Connecticut, Storrs, Connecticut.

§ National Science Foundation, Science Faculty Fellow 1963-1964.

[1] B. Bederson, J. Eisinger, K. Rubin, and A. Salop, Rev. Sci. Instr. 31, 852 (1960).

[2] A. Salop, Ph.D. thesis, New York University, 1961 (unpublished).

[3] A. Salop, E. Pollack, and B. Bederson, Phys. Rev. 124, 1431 (1961).

[4] G. Chamberlain and J. Zorn, Phys. Rev. 129, 677 (1963).

[5] The earlier Scheffers and Stark (Ref. 6) values for lithium and cesium are in error, being about 40% too low for lithium and 20% too low for cesium.

[6] H. Scheffers and J. Stark, Phyzik Z. 35, 625 (1934).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

written (in atomic units) as follows[7]:

$$\alpha_x(m_J)=\alpha_y(m_J)=\sum_{k\neq 0}\left[(J^2+J-m_J^2)\delta_{JJ'}+(J^2+m_J^2+3J+2)\delta_{J+1,J'}+(J^2+m_J^2-J)\delta_{J-1,J'}\right]\frac{|\langle 0|\sum_{i=1}^{N} r_i|k\rangle|^2}{E_k-E_0}, \quad (2a)$$

$$\alpha_z(m_J)=2\sum_{k\neq 0}\left[m_J^2\delta_{JJ'}+\{(J+1)^2-m_J^2\}\delta_{J+1,J'}+(J^2-m_J^2)\delta_{J-1,J'}\right]\frac{|\langle 0|\sum_{i=1}^{N} r_i|k\rangle|^2}{E_k-E_0}, \quad (2b)$$

where J and m_J are the total angular momentum and magnetic quantum numbers of the unperturbed state $|0\rangle$, J' is the total angular momentum quantum number of the state $|k\rangle$, (E_k-E_0) the difference in energy between $|k\rangle$ and $|0\rangle$, r_i the radial coordinate of the ith electron, and N the number of electrons. Since the electric field is parallel to the z axis, the induced dipole is

$$\mathfrak{y}_e=\hat{k}\alpha_z E_z.$$

The resulting force is also in the z direction, and is given by

$$F_e=\alpha_z(m_J)E(\delta E/\delta z), \quad (3)$$

where $\alpha_z(m_J)$ is the "effective" polarizability of the atom in the substate m_J and $\delta E/\delta z$ is the transverse component of the gradient of the magnitude of the electric field. (It is assumed throughout that the nuclear spin is zero.) For substates possessing negative atomic moments, the fields can be adjusted so that the electric and magnetic forces given by Eqs. (1) and (3) are equal and opposite. (In the present experiment these are the positive m_J states.) When this is the case, it follows that

$$\mu_{\rm eff}(m_J)(\delta H/\delta z)=\alpha_z(m_J)E(\delta E/\delta z). \quad (4)$$

Should Eq. (4) apply for several magnetic substates, $\alpha_z(m_J)$ for each substate can be determined separately.

It can readily be shown that the polarizability tensors of all the substates can be uniquely expressed in terms of the effective polarizabilities of the $m_J=+1$ and $m_J=+2$ substates. Thus,

$$\begin{aligned}\alpha_z(0)&=\tfrac{1}{3}[4\alpha_z(+1)-\alpha_z(+2)],\\ \alpha_x(0)&=\tfrac{1}{3}[\alpha_z(+1)+2\alpha_z(+2)],\\ \alpha_x(\pm 1)&=\tfrac{1}{2}[\alpha_z(+1)+\alpha_z(+2)],\\ \alpha_z(\pm 2)&=\alpha_z(+1),\\ \alpha_z(m_J)&=\alpha_y(m_J).\end{aligned} \quad (5)$$

Note that $\bar{\alpha}_x=\bar{\alpha}_y=\bar{\alpha}_z$, where $\bar{\alpha}_x$ is the x component of the polarizability tensor averaged over the five magnetic substates, etc.

[7] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England, 1951). The tensor polarizability is defined on p. 106. Eqs. (2a) and (2b) are obtained by using the T-operator properties tabulated in Eq. (11), p. 63.

Because of the assumed congruency of the electric and magnetic fields, we may write $\delta H/\delta z=kH$ and $\delta E/\delta z=kE$, where k is a geometry factor. Substituting for the gradients in terms of the fields, the balance condition, Eq. (4), can be rewritten as

$$\alpha_z(m_J)=\mu_{\rm eff}(m_J)H/E^2. \quad (6)$$

The measurement of $\alpha_z(m_J)$ is therefore independent of velocity, and specific values of the field gradients and geometry, except in the determination of E from the applied voltage. H can be accurately determined by means of a "zero-moment" experiment,[8] and $\mu_{\rm eff}(H)$ is an accurately known function of H.

Relative measurements, using a normalizing element whose polarizability is known, can also be performed. In the present work, $\alpha_z(1)$ and $\alpha_z(2)$ of metastable $^3P_2$ argon were measured in terms of the polarizability of metastable $^3S_1$ helium, which has been calculated[9] (46.6×10^{-24} cm^3) to an accuracy of about 1%. Substituting for the magnetic field in terms of α(He) in Eq. (6) we obtain

$$\alpha_z(m_J)=\mu_{\rm eff}(m_J)/2\mu_0(V_{\rm He}/V)^2\alpha({\rm He}), \quad (7)$$

where μ_0 is the Bohr magneton, $V_{\rm He}$ is the voltage at balance for metastable helium and V is the voltage balance for the argon substate.

Although the balance condition is valid for beams of infinitesimal width, Salop *et al.*[3] have shown that it is a good approximation for the narrow beams used in the present experiment.

APPARATUS

Figure 1 is a schematic diagram of the apparatus. The metastable beam is formed by effusion of argon into a region where it is cross fired by an electron beam. It then passes through a narrow collimator into the $E-H$ field region, and is finally incident upon a tungsten surface which serves as the cathode of a Bendix electron multiplier. The metastable $^3P_2$ and $^3P_0$ atoms are capable of ejecting electrons from the tungsten cathode, and these electrons are multiplied

[8] V. W. Cohen, Phys. Rev. 46, 713 (1934).

[9] Charles Schwartz (private communication). Also, A. Dalgarno, Proc. Phys. Soc. (London) 72, 1053 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

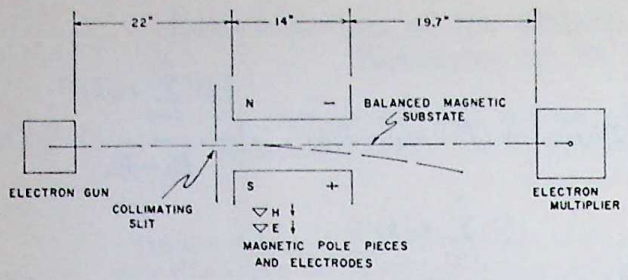


FIG. 1. Schematic diagram of the experiment.

and counted using a preamplifier and scalar. Metastable helium is produced and detected in a similar fashion.

The main chamber contains the magnet pole pieces, which also act as the electrodes for the inhomogeneous electric field. They are insulated from the magnet yoke by a rectangular glass tube which also serves as the vacuum envelope. The electric and magnetic fields between the pole pieces are congruent, since the boundary conditions which they obey are almost identical if the pole pieces are of high permeability and the magnetic fields are sufficiently low to avoid saturation effects.

The pole pieces are 14 in. long, and the ratio of gradient to field is 2.5 cm^{-1} at the beam position, and the gap width is 0.056 in. The beam height as well as its location relative to the pole pieces was fixed by a 0.002-in. $\times$0.050-in. collimating slit mounted directly in front of the pole pieces, at a distance of 0.025 in. from the convex edge. A beam-height limiter 0.060 in. high was mounted at the detector end of the pole piece assembly.

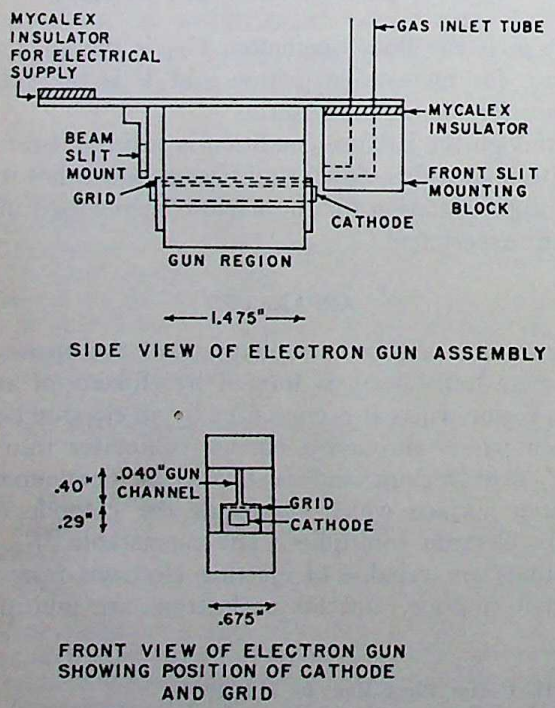


FIG. 2. The electron gun used for metastable production.

A Sensitive Research electrostatic voltmeter was used to measure the applied voltage. Prior to being put into service, its calibration was checked at the National Bureau of Standards. Subsequent calibrations were performed by the manufacturer. The instrument maintained its calibration to within 0.1% throughout the course of the present work.

Figure 2 shows two views of the electron gun.[10]

The gun makes use of a Philips type A cathode, which stands up comparatively well to ion bombardment which results from its operation. The entire gun is mounted between the pole faces of a magnet which produces a field of approximately 1000 G parallel to the electron beam. Recoil broadening was essentially eliminated by the use of a 0.002 in. wide slit mounted at the beam exit of the gun.

The usual operating voltages were 23 V for argon and 26 V for helium, with a current of about 30 mA. At these potentials the photon background was negligible, although it rose rapidly with increasing voltage. For example, at 30 V using argon, about 6% of the detected beam was due to photons. The full beam intensities were typically about 1000 and 200 cpm for argon and helium, respectively. Dark current noise from the multiplier corresponded to only several cpm and so reasonable signal-to-noise was achieved by counting for several minutes, at each pair of E and H field values.

MEASUREMENTS

Measurements were made of the beam intensity as a function of voltage across the electrodes, for a fixed value of magnetic field. Figure 3 presents such a run taken with argon at a field of approximately 260 G. The beam intensity when $V=0$ is due primarily to those atoms with $\mu_{eff}=0$, namely, those in the $^3P_0$ metastable state, and those with $m_J=0$ in the $^3P_2$ metastable state. As voltage is applied across the pole pieces, the beam intensity decreases at first as the atoms without magnetic moment are deflected out by the electric field. The first peak is in balance at about 11 kV and the second peak at 17 kV, and any further increase in voltage results in a monotonically decreasing beam intensity. The curve shape is similar to that of a "zero-moment" curve.[7] Figure 4 shows the region of the single peak for a similar experiment performed with metastable helium at a magnetic field of about 300 G.

Values of the polarizability ratio for the $m_J=+1$ and $m_J=+2$ substates of argon are obtained using a

[10] In the early stages of the experiment an rf discharge was used for metastable production instead of the electron beam. The discharge was pulsed, and a suitably timed gate circuit was used to discriminate the detector against prompt-arrival photons, in favor of the more slowly traveling metastables. Further development of this method was forestalled by satisfactory operation of the electron gun. Nevertheless, the pulse discharge could certainly be improved, and may ultimately prove superior to the gun method in some cases.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

modified form of Eq. (7). Thus,

$$\alpha_z(+1)/\alpha_z(+2)=\tfrac{1}{2}(V_2/V_1)^2, \quad (8)$$

where the subscripts specify the substates. Absolute values, normalized to helium are

$$\alpha_z(+1)=\tfrac{1}{2}(1.5)(V_{\mathrm{He}}/V_1)^2\times 46.6\times 10^{-24}\ \mathrm{cm}^3, \quad (9)$$

$$\alpha_z(+2)=\tfrac{3}{2}(V_{\mathrm{He}}/V_2)^2\times 46.6\times 10^{-24}\ \mathrm{cm}^3. \quad (10)$$

The data are sufficient to determine $\alpha_z(m_J)$ only if the magnetic substates associated with the several peaks can be definitely identified. For example, from the positions of the peak in Fig. 3, a ratio $\alpha_{1z}/\alpha_{2z}\approx 1$ can be inferred if α_{1z} represents the $m_J=+1$ substate or $\alpha_{1z}/\alpha_{2z}\approx\frac{1}{4}$ if α_{1z} represent the $m_J=+2$ substate. (Here, the subscripts 1 and 2 refer to the peaks occurring at the lower and higher voltages, respectively.) It is very unlikely that the ratio should differ greatly from unity,

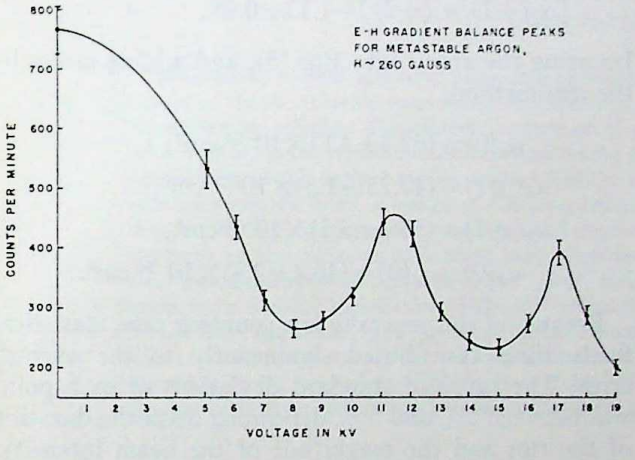


FIG. 3. Beam intensity as a function of applied voltage for metastable argon, at a magnetic field of about 260 G.

since the valence electron, which contributes the principal amount to the polarizability, is in an s state. Nevertheless, a subsidiary experiment was performed in order to definitely identify the substates.

The first peak was brought into balance at a magnetic field of about 260 G and a voltage of 11.1 kV. A beam profile was taken for the balanced substate in steps of 0.001 in. about the beam maximum. The voltage was then dropped to 9.9 kV and a similar profile taken. These two profiles are shown in Fig. 5 The apparatus as used here functioned as a substate selector, and in addition as a relatively crude deflection polarizability spectrometer. A similar deflection experiment was performed on the second balance peak. Beam profiles at balance and at 1.2 kV below balance were taken. Figure 6 shows the results of the second deflection experiment.

An atom experiencing a constant force F transverse to its direction of motion in an atomic beam undergoes a deflection $s=c'F/v^2$, where c' is a constant and v is the

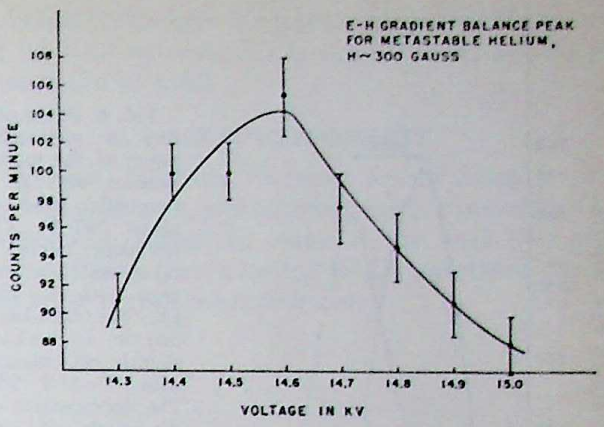


FIG. 4. Beam intensity as a function of applied voltage for metastable helium in the vicinity of the single-balance peak, at a magnetic field of about 300 G.

speed of the atom. In the deflection experiment just described, F is proportional to $\alpha_z(E_0+\delta E)^2-\mu_{\mathrm{eff}}H_0$, where E_0 and H_0 are the balance electric and magnetic fields, and δE, the change in electric field. Dropping the second-order term, F is proportional simply to $\alpha_z E_0\delta E$, or $s=c\alpha_z E_0/v^2$, where c is a constant which includes the δE.

Since δE is the same for the first and second deflection experiments, the ratio of the deflections is given by

$$s_1/s_2=(\alpha_{1z}/\alpha_{2z})(V_1/V_2).$$

Now it can be shown[11] that for deflections small compared to the detector width, a triangular beam is dis-

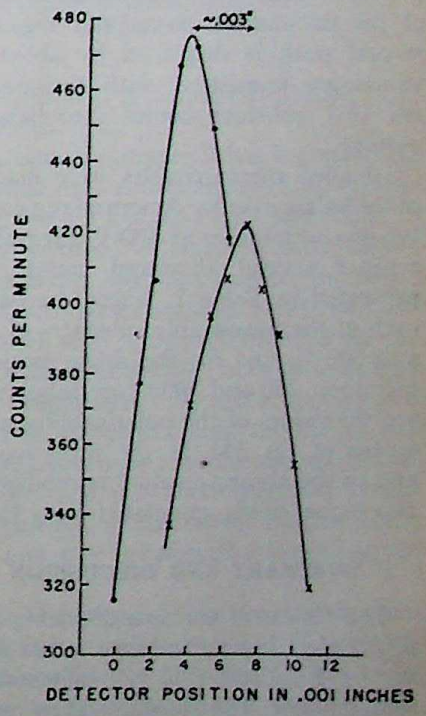


FIG. 5. Beam profiles for metastable argon of the first balance peak, at a magnetic field of about 260 G. The left-hand curve is taken exactly at balance $V=11.1$ kV. The right-hand curve is taken slightly off balance with $V=9.9$ kV. The displacement of the peak of approximately 0.003 in. represents a rough measurement of the polarizability, and therefore serves to identify the balanced substate as $m_J=+1$.

[11] George E. Chamberlain, Ph.D. thesis, Yale University, 1961, p. 83 ff. (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

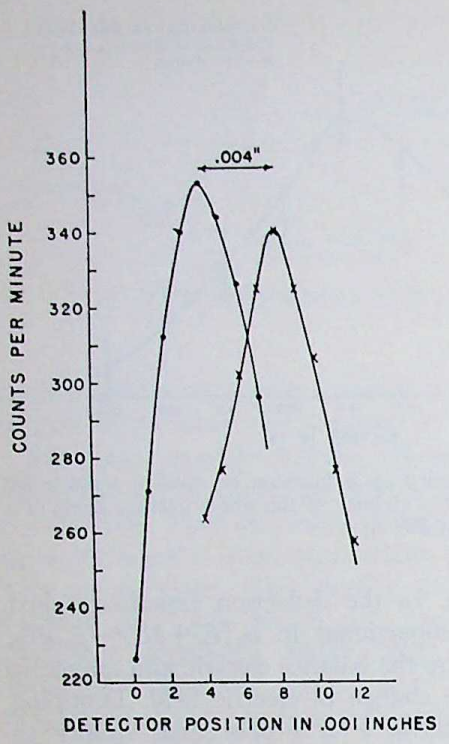


FIG. 6. Beam profiles for metastable argon of the second balance peak at a magnetic field of about 260 G. The left-hand curve is taken exactly at balance with $V=16.9$ kV. The right-hand curve is taken slightly off balance, with $V=15.7$ kV. The displacement of the peak of approximately 0.004 in. represents a rough measurement of the polarizability and therefore serves to identify the balanced substate as $m_J=+2$.

placed with an unaltered profile. That is, velocity broadening of the beam shape is negligible. Our deflections are actually comparable to the beam width, so that some velocity broadening should occur. Nevertheless, the beam shape will be only slightly altered, and it should be possible to determine whether the ratio is of the order of unity or four. Figure 5 shows that the first peak is displaced by approximately 0.003 in. as a result of the unbalanced force, and Fig. 6 shows that the second peak is displaced by about 0.004 in. These values are consistent with the assumption that the $m_J=+1$ substate comes into balance at the lower voltage.

Detailed measurements were made in the vicinity of the balance peaks. A normal run consisted of counting for several minutes at 100-V intervals, usually so as to register several thousand counts. The results are presented in Table I, where the balance voltages for each of the measurable substates of argon at 200, 260, and 300 G, and for the single measurable substate of helium at 260 and 300 G are tabulated. Also tabulated are the values of the polarizability ratios calculated by means of Eq. (8), at the three magnetic fields used, and of the absolute values normalized to helium at the two higher fields, calculated using Eqs. (9) and (10).

SUMMARY AND DISCUSSION OF RESULTS

The results of our measurements of the z component of the electric polarizability tensor of the $m_J=+1$ and $m_J=+2$ substates of $^3P_2$ metastable argon are presented below. The absolute values represent a combined average of the results taken at magnetic fields of 260 and 300 G. The ratio measurement represents a

TABLE I. Summary of experimental results.

Approximate magnetic field (gauss)	Balance voltage (kV)	Magnetic substate and element	$\alpha_z(m_J)$	$\alpha_z(+1)/\alpha_z(+2)$
200	9.9	+1(Ar)		1.13
200	14.9	+2(Ar)		
260	11.2	+1(Ar)	51.5	1.14
260	16.9	+2(Ar)	45.3	
260	13.6	+1(He)		
300	12.3	+1(Ar)	49.3	1.13
300	18.5	+2(Ar)	43.8	
300	14.6	+1(He)	(46.6)[a]	

[a] Calculated by Schwartz and Dalgarno (see Ref. 8).

combined average of the results taken at magnetic fields of 200, 260, and 300 G. We find

$$\alpha_z(+1)=(50.4\pm3.5)\times10^{-24}\ \text{cm}^3,$$
$$\alpha_z(+2)=(44.5\pm3.1)\times10^{-24}\ \text{cm}^3,$$
$$[\alpha_z(+1)/\alpha_z(+2)]=1.13\pm0.08.$$

Inserting the above into Eq. (5), and adding errors by the rms method,

$$\alpha_z(0)=(52.4\pm4.8)\times10^{-24}\ \text{cm}^3,$$
$$\alpha_x(\pm1)=(47.25\pm2.3)\times10^{-24}\ \text{cm}^3,$$
$$\alpha_x(\pm2)=(50.4\pm3.5)\times10^{-24}\ \text{cm}^3,$$
$$\alpha_x(0)=\alpha_y(0)=(46.4\pm2.4)\times10^{-24}\ \text{cm}^3.$$

Because of the generally low counting rate, statistical fluctuations contributed significantly to the over-all error. The range of standard deviations of each point was between 1% and 3%, depending upon the duration of the run and the magnitude of the beam intensity, with the helium data usually having the larger deviations. However, the argon peaks were somewhat broader than the helium peaks, owing to the greater overlapping of the three substates $m_J=+2, +1, 0$ of $^3P_2$ argon, combined with the presence of the single substate of $^3P_0$ argon, as compared to the presence of the two substates $m_J=+1, 0$ of the $^3S_1$ state of helium, combined with the presence of the single substate of $^1S_0$.

Among the other causes of experimental error are the following: an apparent shift in the balance peak due to fast atoms from unbalanced substates; fluctuations of electron gun current, source pressure, and magnetic and electric fields during the course of a run; measurement of the voltage using the NBS calibrated voltmeter; lack of perfect parallelism of the beam with respect to the pole face; and deviations from the balance condition over the finite beam width. A detailed discussion of most of these errors is contained in Ref. 3.

We make an over-all assignment of $\pm7\%$ to the error in measuring $\alpha_z(+1)$ and $\alpha_z(+2)$, based upon conservative estimates of the above error sources.

It was to be expected that the polarizabilities of

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$^3P_2$ argon would be quite large, since its structure is grossly similar to that of ground-state potassium, which has a polarizability of about 36×10^{-24} cm^3. Further, since the valence electron contributes the major amount to the polarizability, the polarizability ratio $\alpha_z(+1)/\alpha_z(+2)$ should be quite close to unity. (The valence electron is in an s state.)

The P configuration of the atom is due to the five p electrons of the incomplete M shell, whose contribution to the polarizability should therefore contain a strong anisotropy. However, since the contribution of the core to the total polarizability is small, the net anisotropy should also be small.

ACKNOWLEDGMENTS

We wish to thank Professor Arthur Salop for his continued interest, and for many useful discussions and suggestions during the course of the work. We also wish to thank Judah Levine for his assistance during the latter part of the experiment.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Studies of the Mechanism of Electron-Ion Recombination. II

WILLIAM A. ROGERS* AND MANFRED A. BIONDI†
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania
(Received 13 January 1964)

Microwave, optical spectrometric, and interferometric techniques have been used to investigate the nature of the electron-ion recombination process in low-pressure, pure helium afterglows at 300 and 77°K. To determine whether dissociative capture of electrons by helium molecular ions, i.e., $He_2^+ + e \rightarrow He^* + He + (KE)$, is the recombination reaction taking place, we have attempted to detect the kinetic energy of dissociation of the excited atoms produced in the recombination process by a study of the widths of the emitted afterglow lines. A model of the low-pressure helium afterglow which reasonably accounts for our electron density, line intensity, and linewidth observations involves principally the creation of He^+ ions and electrons by the collisions of pairs of helium metastable atoms, the three-body conversion of He^+ ions to He_2^+ ions, the ambipolar diffusion of He^+, He_2^+, and electrons, and the diffusion of helium metastable atoms to the walls. The λ5876 radiation (3^3D-2^3P transition) from the later part of the afterglow, when recombination of He_2^+ and electrons is expected to be the source of the excited atoms, exhibits a substantial broadening over the approximately thermal width observed during the discharge and in the early afterglow, when the radiation is expected to originate from electron impact excitation of helium atoms. It is concluded that the dissociative recombination process is the likely source of the observed afterglow line broadening.

I. INTRODUCTION

NUMEROUS experimental studies of the afterglows of pulsed microwave discharges[1–4] have shown that volume electron-ion recombination appears to be an important or dominating process in the removal of free electrons from ionized gases for several different gases. The principal technique in these experiments has involved the determination of spatially averaged electron densities during the afterglow by means of microwave probing techniques (see part I of this paper,[5] hereinafter referred to as I) and an analysis of the time dependence in terms of solutions of the equation of continuity for electrons subject to losses by recombination and diffusion. This analysis, as commonly employed, is outlined in Eqs. (3)–(6) of I.

In the rare gases neon and argon, it has been possible to obtain experimental data consistent with the recombination solution, $1/n_e=[1/n_e(0)]+\alpha t$, and thus to evaluate the recombination coefficient α with reasonable precision. In helium, however, due to the larger ratio of diffusion loss to recombination loss it has not been possible to obtain such a clear distinction. In addition, the helium measurements have been complicated by the fact that at pressures of only a few mm Hg the atomic spectra, which are presumably related to the recombination process, give way to molecular spectra of excited He_2. Thus, serious doubts have been cast on the validity of measurements of the recombination coefficient in helium by means of the afterglow time dependence studies.[6]

By means of a set of controlled mixture experiments in helium and argon, it has been shown (see I) that the large recombination coefficients observed in the pure gases do not appear under circumstances that inhibit

---

* Present address: Physics Department, Thiel College, Greenville, Pennsylvania.
† Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania.

[1] M. A. Biondi and S. C. Brown, Phys. Rev. 76, 1697 (1949).
[2] R. B. Holt, J. M. Richardson, B. Howland, and B. T. McClure, Phys. Rev. 77, 239 (1950).
[3] M. A. Biondi and T. Holstein, Phys. Rev. 82, 962 (1951); M. A. Biondi, *ibid.* 83, 1078 (1951).
[4] A. C. Faire and K. S. W. Champion, Phys. Rev. 113, 1 (1959).
[5] M. A. Biondi, Phys. Rev. 129, 1181 (1963).
[6] C. S. Leffel, M. N. Hirsch, and D. E. Kerr, Ann. Phys. (N. Y.) (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the formation of diatomic molecular ions, thus supporting the hypothesis of Bates[7] that the process under observation is the dissociative recombination of molecular ions with electrons in the reaction,

$$X_2^+ + e \rightarrow X^* + X + \text{kinetic energy}. \quad (1)$$

This naturally leads to the supposition that the same process is responsible for any large recombination loss occurring in moderate-pressure helium afterglows.

The purpose of the experiments reported in this paper has been (a) to establish definitely the presence or absence of the dissociative recombination process in pure helium afterglows at low pressures (2 mm Hg or less) by a technique which avoids the complexity and uncertainty of interpretation inherent in the microwave electron-density determinations under mixed diffusion and recombination conditions, and (b) to provide an estimate of the magnitude of the recombination coefficient, if possible. The principal new technique introduced is the interferometric measurement of the intensity profile of an optical emission line of the helium atomic spectrum.

The significance of this measurement can be understood by reference to Fig. 1 and to Eq. (1) and from the following considerations. The dissociative process can be efficient only if it is exothermic, i.e., the transition from the bound state A to the dissociated state B must be accompanied by the conversion of potential energy into kinetic energy in a significant amount. This converted energy is shown as the vertical scaled length E_D in the figure. The kinetic energy so produced is shared by the fragments which fly out from the dissociating system. In the case of a homonuclear diatomic molecule, this energy of dissociation E_D is shared equally by the two identical atoms. Also, in the case of helium, the relative energies of the molecular-ion ground state and the molecular and atomic states are such that one of the atoms will almost certainly be produced in an electronically excited state.

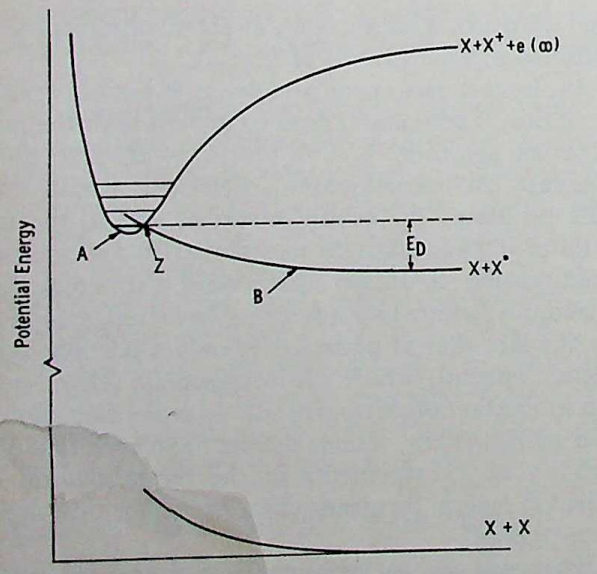


FIG. 1. Schematic representation of the interatomic potential curves of interest in the dissociative recombination of X_2^+ ions with electrons. The upper potential curve is that for a molecular ion and an electron at rest at infinity.

[7] D. R. Bates, Phys. Rev. 77, 718 (1950); 78, 492 (1950); 82, 103 (1951).

If this excited atom is produced in a radiating state of sufficiently short lifetime, it can be expected to radiate some of its excitation energy before it has time either to lose its kinetic energy through collisions or to lose its excitation energy by resonant transfer to a slowly moving atom. If the atom does radiate before experiencing energy loss of either kind, the emitted photon will be shifted from its normal wavelength by the Doppler effect due to the atomic velocity. Since there is no preferred direction for dissociation, the velocities of the dissociating atoms will be randomly directed, and consequently, the observed effect will be a Doppler broadening of the atomic spectral line in excess of the broadening due to the temperature of the molecular ions by an amount corresponding to the kinetic energy $E_D/2$. The direct observation of such anomalous line broadening may be taken as positive evidence of the predominance of the dissociative recombination process in the production of the afterglow emission spectrum.

For this type of measurement helium has been chosen, rather than neon or argon, primarily because of the simplicity of its optical spectrum and the short lifetimes of some of the radiating states. Since the analysis involves considerable computation, a single line (λ5876, corresponding to the transition $3\,^3D$–$2\,^3P$) has been chosen for study.

The linewidth studies have been correlated with electron-density measurements and with time-resolved emission spectra. In addition to the recombination study itself, attention has been given to the early part of the afterglow preceding the "recombination" period which has been the subject of most previous investigations. An attempt has been made to formulate a complete model of the afterglow processes for a simple case.

II. EXPERIMENTAL SYSTEM

The complete experimental arrangement is shown diagrammatically in Fig. 2, except for the ultrahigh-vaccum gas handling system. The apparatus shown may be regarded as comprising three main units: microwave, spectrometer, and interferometer systems.

A. Microwave System

At the center of the figure is the microwave resonant cavity which encloses the discharge vessel and has hollow walls which may be filled with liquid nitrogen. A repeated, pulsed discharge is produced in the quartz discharge vessel by power from the magnetron. Following this pulse a very low power probing signal is applied from the klystron. Power reflected from the cavity

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

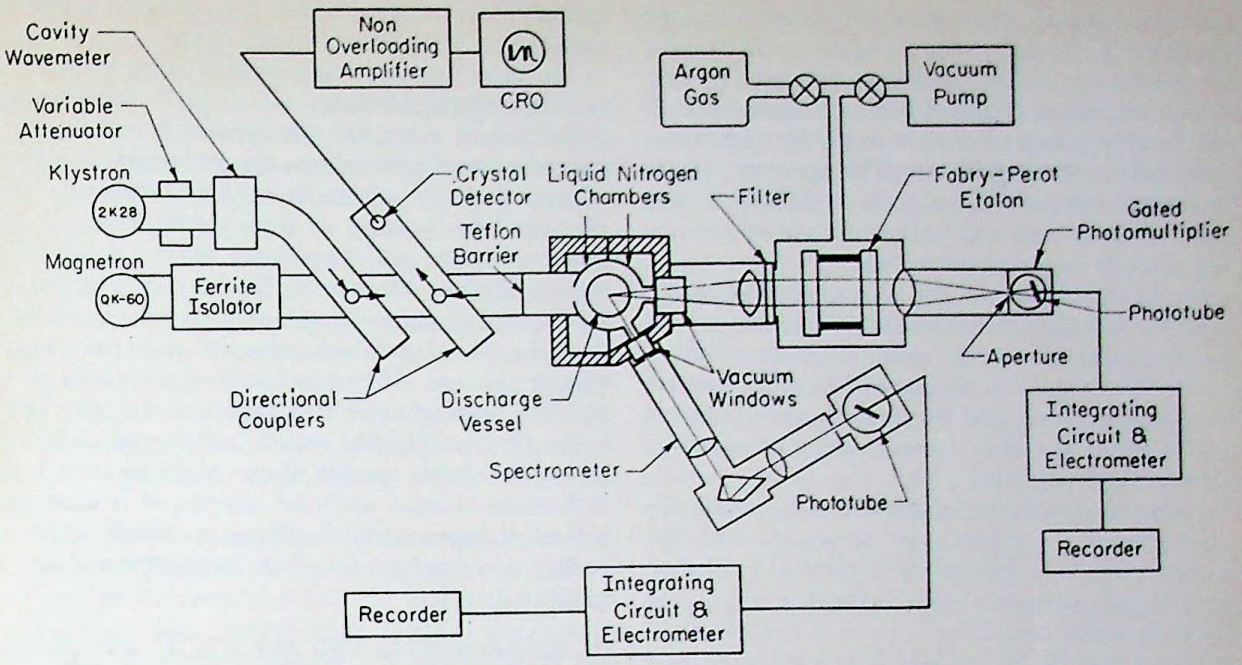


FIG. 2. Block diagram of the combined microwave, optical spectrometric, and optical interferometric systems used to study afterglows.

resonator is detected by the crystal detector and passed through an amplifier, which suppresses the strong magnetron pulse. The resulting signal is displayed on a synchronized oscilloscope.

B. Spectrometer System

A spectrometer port is provided by a narrow slit which tunnels through the double wall of the cavity resonator and an evacuated double glass window which tunnels through the foam insulation. The spectrometer has a synchronous prism table drive motor and an electronically gated multiplier phototube detector, whose output goes to an integrating filter, dc electrometer and synchronously driven strip chart recorder.

C. Interferometer System

A second small opening in the cavity wall, similar to the spectrometer port but circular in cross section, provides a port for the Fabry-Perot etalon. The etalon and detection system used here[8] is similar in principle to that used by Jacquinot and Dufour.[9] It is illustrated in more detail in Fig. 3.

The mechanical support and adjustment of the etalon plates (F) is provided by means of a conventional arrangement of invar spacers (G) and spring mounted compression pins (H) similar to those described in textbooks,[10] with modifications as required to permit operation through the walls of the evacuated housing (C). The photoelectric detection, integrating, and recording system is identical with that of the spectrometer system.

In use the etalon, Fig. 3, is placed so that the focal plane of the collimating lens (A) is near the center of the microwave cavity, i.e., in the middle of the discharge. The collimated beam passes from the lens through a glass color filter and multilayer interference filter (B),

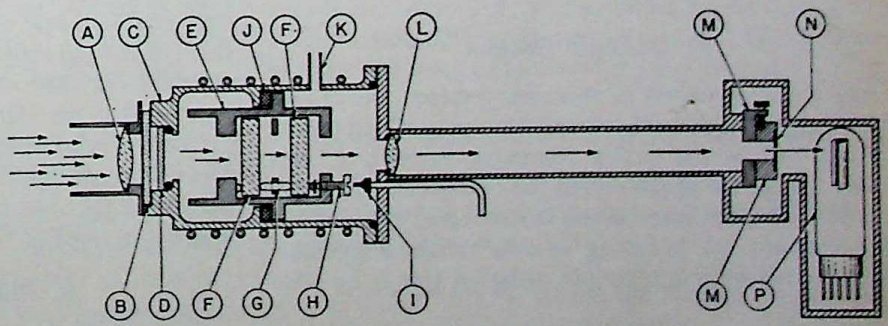


FIG. 3. Detailed cross section view of the photoelectric recording, Fabry-Perot interferometer. (See text for explanation.)

[8] M. A. Biondi, Rev. Sci. Instr. 27, 36 (1956).
[9] P. Jacquinot and C. Dufour, J. Rech. Centre Natl. Rech. Sci. Lab., Bellevue (Paris) 6, 91 (1949).
[10] See, for example, S. Tolansky, *High Resolution Spectroscopy* (Methuen and Co. Ltd., London, 1947).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

through the optically clear vacuum window (D), the etalon plates (F), and the vacuum window and focusing lens (L). Here it is focused on an opaque disk (N) having in it a circular aperture of 0.051 cm diam which is centered on the interference pattern by the micrometer carriages (M) to within 0.001 cm. This aperture permits a very limited angular range of incident light, including a portion of the central spot of the interference pattern, to pass through to the photocathode.

From simple considerations of geometrical optics, one can calculate the luminous flux reaching the cathode of the multiplier phototube, assuming lossless optical elements. This calculation, together with rough measures of the absorption losses and the photocathode quantum efficiency, forms the basis for estimating absolute intensities in the afterglow.

The observed resolution of the etalon in practice is two parts in 10^6, in approximate agreement with the calculated value. The experimental value of resolution is based on measurements made (a) with a microwave discharge in argon containing a trace of isotopic Hg^{202} using the narrow Hg line at 5461 Å as a standard, and (b) on the assumption that the smallest attainable values of linewidth in helium are due solely to thermal Doppler broadening corresponding to the ambient temperature. The values deduced by the two different methods are in good agreement.

A wavelength scanning system with linear time base is provided for the etalon by continuously increasing the refractive index μ in the space between the etalon plates (see Fig. 2). This is achieved by first evacuating the interior of the sealed etalon housing with a vacuum pump to provide a starting point and then admitting gas through a preset needle valve connected to a source of argon at 200 atm pressure. The gas density in the etalon rises slowly from zero to slightly over one atmosphere, at which point further increase is interrupted by the opening of a safety valve. This system offers a uniform scanning rate over approximately 10 fringes. Since the synchronous motor drive on the recorder provides a linear time base for the chart paper, one obtains a linear presentation of intensity as a function of frequency (wavelength) within each interference order, thus giving us the desired line shapes directly.

D. Vacuum System

The gas is admitted to and removed from the quartz discharge vessel through a small bore tube which passes directly out of the cavity resonator through a close fitting hole at the upper end of the cavity axis. Through this tube the discharge vessel is connected to an ultrahigh-vacuum gas handling system[11] with a bakeable oil manometer[12] and a copper isolation trap.[13] The discharge vessel itself is isolated from the rest of the vacuum system by a liquid-nitrogen trap.

The helium gas for these experiments is prepared by distillation from the liquid into 1-l gas flasks. The flasks, connected to a special glass manifold which can be opened under liquid helium, are first prepared by being outgassed under ultrahigh-vacuum conditions; then the manifold is sealed off from the outgassing system under vacuum and transferred to the liquid-helium storage Dewar. After being filled by distillation from the liquid helium, the flasks are individually sealed off from the special manifold and sealed on to the operating vacuum system. After a bake-out of the whole system at 400°C overnight and the attainment of a vacuum level below 10^{-9} mm Hg, the sealed flask is opened. By comparison of visible spectra of our prepared pure helium with those of mass analyzed samples of reagent grade helium, it appears that in all gas samples used the impurity level has been less than one part in 10^7 and too small to detect.

III. EXPERIMENTAL AND ANALYTICAL PROCEDURES

A. Microwave Procedures

The average electron density as a function of time is determined by the technique of microwave reflection from the cavity resonator. This technique has been discussed in detail elsewhere and will not be reviewed here.[14]

In view of the high Q of the cavity resonator, it is not surprising that the growth of the discharge in response to the applied power should be quite complicated. The initial conditions of the afterglow seem to be most sensitive to the length of the power pulse and less sensitive to the repetition rate and power delivered from the magnetron. With the magnetron frequency set slightly above the resonant frequency of the cavity without discharge, the buildup of plasma density during the power pulse shifts the cavity resonance through the magnetron frequency to a higher frequency, at which point further increase in plasma density is prevented by the self-detuning of the cavity toward still higher frequencies. Observations have shown that in this system the discharge has passed through a maximum of light output and settled down to approximately the steady-state emission rate after about 40 μsec. Therefore, this value of the pulse duration has been set as an arbitrary standard for most of the experiments. Experience also showed that an afterglow period of about 2 msec is the least time in which the essential features of the time-dependent afterglow processes can be exhibited and, therefore, this has been used as the pulse repetition interval in most of the experiments.

B. Spectrometric Procedures

Using the high-speed optical spectrometer with gated multiplier phototube and recorder, comparisons have

[11] D. Alpert, J. Appl. Phys. 24, 860 (1953); *Handbuch der Physik* (Springer-Verlag, Berlin, 1957), Vol. 12.
[12] M. A. Biondi, Rev. Sci. Instr. 24, 989 (1953).
[13] D. Alpert, Rev. Sci. Instr. 24, 1004 (1953).
[14] M. A. Biondi, Rev. Sci. Instr. 22, 500 (1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

been made of the intensities of several atomic lines at various times in the afterglow. The synchronous motor drives on the chart recorder and the spectrometer prism table provide a direct conversion of wavelength to chart length scale. By setting the electronic gate to any desired duration and to any desired onset delay relative to the end of the discharge, the afterglow spectrum has been recorded during each of several short periods (durations of 50–500 μsec) in the afterglow. Also, by stopping the prism table drive on the peak of any desired line and periodically changing the onset delay time of the gate, the time dependences of several individual lines have been traced. Attention was centered mainly on the diffuse series lines λ4472 and λ4922 ($4D$-$2P$) and λ5876 and λ6678 ($3D$-$2P$), although other lines of the sharp, principal, and diffuse series were also monitored.

Even at the low pressures of these experiments the spectrometer indicates the presence of some helium molecular-band intensity, which the etalon cannot resolve but shows as a "white" background to the desired atomic line emission. Therefore, the spectrometer has been used to determine the relative strengths of the bands that fall within the transmission range of the multilayer interference filter used with the etalon. The transmission of the interference filter together with its associated color filter was determined by comparing the spectrum of an incandescent lamp with and without interposition of the filter. Although the spectrometer does not resolve detailed structure of the bands and records atomic lines only as impulses whose shapes are fixed by the time constant of the electronic integrating circuit, the relative total strength (intensity integrated over wavelength range) of a band to a line is given by the ratio of the areas under the respective recorded contours. This information has been used to correct for the apparent line broadening which would result from the superposition of the band radiation on that of the line recorded by the etalon.

C. Linewidth Measurements with the Etalon

Several factors enter into the production of the observed atomic linewidth. Some of these factors are analytically separable; some must be treated as they appear. We now outline the analysis in terms of which the line broadening measurements are interpreted.

1. Single Component Line Shape due to Temperature and Dissociative Recombination

We assume that line broadening is entirely due to Doppler effects arising from temperature motion and from dissociation of the excited molecule. The kinetic-energy distribution of the diatomic ions is assumed to be Maxwellian at the gas temperature, so that it may be written

$$F_0(v)=(\beta/\pi)^{3/2}\exp(-\beta v^2), \qquad (2)$$

where $\beta=M/kT$, M being the mass of a single helium atom. The problem is further simplified by assuming that all dissociations leading to a given excited atomic state originate in a single specific molecular configuration, and, therefore, release a definite total kinetic energy E_D which is shared equally by the two dissociating atoms. Thus, the speed of the excited atom relative to the center of mass of the diatomic ions is $v_1=(E_D/M)^{1/2}$. The direction taken by the atom is assumed to be randomly oriented with respect to the observer, so the distribution F_1 in velocity relative to the center of mass is simply a unit impulse function (δ function) of speed at v_1.

For the computation of Doppler shift we are interested only in the distribution in speed along the observer's line of sight (the z axis). For the thermal distribution we get by integration over the x-y plane

$$F_0(v_z)=(\beta/\pi)^{1/2}\exp(-\beta v_z^2), \qquad (3)$$

while the integration of F_1 over the x-y plane yields

$$F_1(v_z)=1/2v_1 \qquad (4)$$

for $-v_1\leqslant v_z\leqslant v_1$ and zero elsewhere.

The two distributions F_0 and F_1 are combined in the following way. For every v_z' at which atoms are repelled from the center of mass there is a complete Maxwellian distribution of velocities of the center of mass, giving rise to a group of atoms having a Maxwellian distribution centered on the value v_z'. That is, the number of atoms in the speed range dv_z' at v_z-v_z' is

$$dG=F_0(v_z-v_z')F_1(v_z')dv_z'. \qquad (5)$$

The total resultant distribution in v_z is then obtained by integrating Eq. (5) over the entire range of v_z'. If we first substitute for v everywhere the corresponding Doppler wave number shift, $\tilde{\nu}=\tilde{\nu}_0 v/c$, and simplify notation by putting $a=(Mc^2/kT)^{1/2}/\nu_0$ and $b=(E_D/kT)^{1/2}$, we obtain as the resultant intensity contour (normalized to unit area)

$$G(\tilde{\nu})=(a/4b)[\mathrm{erf}(a\tilde{\nu}+b)-\mathrm{erf}(a\tilde{\nu}-b)]. \qquad (6)$$

Analysis of $G(\tilde{\nu})$ shows that at the center of the line ($\tilde{\nu}$ much smaller than $\tilde{\nu}_1\equiv\tilde{\nu}_0 v_1/c$) the slope of $G(\tilde{\nu})$ equals the slope of $F_0(\tilde{\nu})$ multiplied by $\exp(-E_D/kT)$. If E_D is reasonably large compared to kT, the line shape $G(\tilde{\nu})$ will be tend to become very flat topped.

2. Instrumental Broadening of the Single Component Line

The fringe shape for monochromatic light of wave number $\tilde{\nu}$ passing through an etalon at normal incidence is given by Airy's formula[10]

$$I=T^2(1-R)^{-2}[1+K\sin^2(2\pi\mu l\tilde{\nu})]^{-1}, \qquad (7)$$

where T is the transmission and R the reflectivity of the coatings on the etalon plates, μ is the index of refraction of the medium between the plates, l the plate

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

spacing, and $K=4R/(1-R)^2$. For present purposes we will omit the intensity reduction factor $T^2(1-R)^{-2}$ and write for the normalized etalon transmission function,

$$H(f)=(1+K\sin^2\pi f)^{-1}, \tag{8}$$

in which $f=2\mu t\bar{\nu}$. The quantity f is the interference order number; it is used here as a dimensionless frequency variable measured in units of the free spectral range of the etalon. This is a simple scale change, which does not alter the form of G, Eq. (6), and provides a convenient unit for discussion of the etalon measurements.

The line shape resulting from the transmission of the line $G(f)$ through the etalon is obtained by considering the line as the superposition of fringes of all the monochromatic elements whose integral over frequency constitutes $G(f)$. Thus, if the etalon is set or "tuned" for a frequency f (i.e., the interference order is integral for f), then light at frequency f' produces a fringe intensity contribution of $G(f')H(f-f')df'$. The total resultant intensity transmitted at a given etalon setting is then

$$\phi(f)=\int_0^\infty G(f')H(f-f')df'. \tag{9}$$

This integral requires numerical computation. It has been computed at intervals of $f=0.01$ for several values of T and E_D, initially by manual tabulation and then more extensively by the computer facility of the Mathematics Department of the Westinghouse Research Laboratories. From the results so obtained, it was evident that no simple procedure would satisfactorily separate the instrumental contribution from the actual line broadening. For this reason the computed over-all profiles are compared directly with the observed total linewidths.

3. *Choice of the λ5876 Line and the Shape of the Triplet Line in the Etalon*

Although the singlet lines offer the obvious advantage of simplicity in their lack of fine structure, the λ5876 ($3\,^3D-2\,^3P$) line was selected because it has the shortest lifetime among all the visible transitions ($\tau_{\text{rad}}=1.34\times10^{-8}$ sec), is well isolated from neighboring lines, is known to be a bright line in low-pressure afterglows, and is in a wavelength region suited to available photoelectric tubes. Moreover, its terminal state, $2\,^3P$, is itself a radiating state and, therefore, not sufficiently populated to cause any spurious broadening by self-absorption. The high radiative transition rate of nearly 10^8 sec^{-1} was initially considered necessary to assure a high probability of radiation prior to excitation transfer or kinetic energy loss by collisions, which have been estimated to occur at a rate of the order of 10^8 per second at a pressure of one mm Hg. Thus, measurements of line shapes have been carried out at pressures of a few mm Hg or less.

An analysis of the structure of the λ5876 line had been made by Brochard *et al.*[15] using an etalon and a helium source cooled in liquid helium. Their results have been adopted as a basis for the construction of the expected line shapes. In our experiments it is not possible to separate each of the six components of the line because of temperature broadening. The etalon-free spectral range has been chosen to make the one separable component appear as a satellite one and one-half orders removed from the main peak composed of the other five. The fringe contributions of all six have been added graphically over the range of interest in the main peak, and resulting widths at half-intensity have been measured from the composite graph for a variety of parameter values. The half-widths of these lines have been compared with actually observed half-widths to determine E_D.

4. *Spurious Line Broadening Effects*

In addition to the temperature and dissociative recombination line broadening effects, there are two principal spurious effects which may lead to an apparent broadening of the observed lines. These are (a) the presence of molecular band radiation in the light reaching the photocathode, and (b) the effect of electrical integrating networks between the phototube and the electrometer.

Any molecular band radiation which passes through the interference filter and the etalon may be assumed to consist of numerous rotational transition lines, a fact which is corroborated by the shapes of the molecular bands as recorded through the spectrometer. These rotation lines are spaced nonuniformly in frequency and irrationally with the etalon-free spectral range. Consequently, they overlap each other in a random fashion and form what may be regarded as "white" light. To compare the effect of band radiation with that of an isolated atomic line we make the simplifying assumption that the etalon transmission function $H(f)$ is periodic in f with unit period, is unity over a range f_e equal to the etalon half-width, and zero in the remainder of each unit range; that the atomic line intensity $G(f)$ is a constant A over a range f_a equal to the half-width of the line and zero elsewhere, f_a being larger than f_e but less than unity; and that the molecular band intensity is a constant B over a frequency range f_b which is much greater than unity and zero elsewhere. With these very simple line shapes it can be readily shown that the etalon reduces the ratio of total band radiation to line radiation by the factor f_a. Checks on this procedure by graphical area measurements on the carefully computed etalon-Doppler lines show that the simple theory is always conservative in that it overestimates the molecular background by about 10–30%. The effect of a white background on atomic line half-width measurements is to displace the apparent zero intensity reference of the line downward, causing the half-width

[15] J. Brochard *et al.*, J. Phys. Rad. 13, 433 (1952).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to be measured too low on the line profile, thereby giving too large a value for the half-width. A correction for this effect has been made which is accurate for narrow lines and/or small "white" backgrounds and which overestimates the spurious broadening for wider lines and/or larger "white" backgrounds. Thus, the deduced linewidths in the presence of molecular light background are smaller than the actual widths.

The electrical integrating circuits necessary to suppress noise and to convert the time-gated sample signals into dc signals for amplification and recording place a limit on the speed with which the line profile may be scanned. In practice, if the circuits are characterized by a response time constant τ and the line profile is traced between half-intensity points in a time Δt, care is taken to assure that $\Delta t \gtrsim 5\tau$ for the narrowest line in a set of measurements, leading to a distortion of the recorded profile which increases the apparent line breadth by less than 2%. The error in measured width for broader lines, which take even longer to trace, is even smaller.

IV. MODEL OF THE AFTERGLOW

The variations of the measured electron densities, spectral intensities, and line shapes during the afterglow are sufficiently complicated that it is of considerable help to have in mind the relevant atomic-collision processes which contribute to their behavior. The following model of the afterglow is consistent with previous researches on helium[16–18] and with the observations obtained in the present studies.

At the low pressures (~1 mm Hg) and moderate container size (fundamental diffusion length, $\Lambda_1 = 0.735$ cm) used in these afterglow studies, the principal loss of the electrons, ions, and metastable atoms is by diffusion to the container walls. In addition, there is substantial production of electrons and atomic ions by collisions of pairs of metastable atoms,[19]

$$\mathrm{He}^M + \mathrm{He}^M \rightarrow \mathrm{He}^+ + e + \mathrm{He}. \qquad (10)$$

Since the excitation energy of the metastable state of the helium atom is ~20 eV and the ionization potential ~25 eV, the electrons produced by this reaction have ~15 eV kinetic energy. As will be discussed more fully later, it is to be expected that before this excess kinetic energy is completely dissipated by elastic collisions, there is a substantial probability that these electrons will collide with metastable atoms, raising them to higher, radiating states, i.e.,

$$\mathrm{He}^M + e(\mathrm{fast}) \rightarrow \mathrm{He}^* + e(\mathrm{slow}). \qquad (11)$$

[16] M. A. Biondi and S. C. Brown, Phys. Rev. 75, 1700 (1949).
[17] A. V. Phelps and S. C. Brown, Phys. Rev. 86, 102 (1952).
[18] A. V. Phelps and J. Molnar, Phys. Rev. 89, 1202 (1953); A. V. Phelps, *ibid.* 99, 1307 (1955).
[19] M. A. Biondi, Phys. Rev. 88, 660 (1952). Contrary to the conclusion of H. Myers, Phys. Rev. 130, 1639 (1963), available experimental evidence (Refs. 17–19) strongly supports the conclusion that metastable-metastable ionizing collisions lead to the production of *atomic*, rather than molecular, helium ions.

The atomic ions produced in the ionizing reaction, Eq. (10), are converted to molecular ions in three-body collisions,[17]

$$\mathrm{He}^+ + 2\mathrm{He} \rightarrow \mathrm{He}_2^+ + \mathrm{He}. \qquad (12)$$

While the major loss of molecular ions is by diffusion to the walls, dissociative recombination between molecular ions and electrons is expected to produce radiating atomic excited states according to the reaction

$$\mathrm{He}_2^+ + e \rightarrow \mathrm{He}^* + \mathrm{He}. \qquad (13)$$

If we neglect small terms in the equations governing the production and loss of the atomic ions and the molecular ions, we obtain the following approximate relationships:

$$\partial n_1/\partial t \simeq D_{a1}\nabla^2 n_1 - K_c n^2 n_1 + \beta_i n_m^2 \qquad (14)$$

and

$$\partial n_2/\partial t \simeq D_{a2}\nabla^2 n_2 + K_c n^2 n_1, \qquad (15)$$

where D_{a1} and D_{a2} are the ambipolar diffusion coefficients of the atomic and molecular ions of densities n_1 and n_2, respectively, n is the neutral atom density, n_m the metastable concentration, K_c the three-body conversion coefficient for reaction (12), and β_i is the two-body ionizing coefficient for metastable-metastable collisions, Eq. (10). It is assumed that the above equations apply to thermal energy particles, since the electron- and ion-energy decay times are short ($\lesssim 10$ μsec) compared to the time scale of the measurements.

Since the molecular-ion concentration remains small compared to the atomic-ion density throughout these low-pressure afterglows, we may take the electron density, $n_e \simeq n_1$. Also, at low pressures (<1 mm Hg) the metastable atoms decay principally by diffusion to the walls; thus, the third term on the right of Eq. (14) may be approximated by $\beta_i n_m^2(0) \exp(-2\nu_m t)$, where ν_m is the diffusion frequency of the metastables. From the published values[16–18] of the atomic ion, molecular ion, and metastable atom diffusion coefficients, of K_c and β_i (see Table I), and assuming a metastable atom concentration sufficient to produce the observed maximum value for the electron density in the afterglow, i.e., $n_m(0) \sim 10^{12}$ cm^{-3}, one can calculate the afterglow behavior of the atomic ions, electrons, and molecular ions. The results of a model afterglow calculation at a pressure of 0.54 mm Hg and $T = 300$°K are shown in Fig. 4. Starting from a very small number of electrons

TABLE I. Coefficients used in the afterglow analysis (for references, see text).

		He^+	He_2^+	He^M ($2\,^3S$)
Diffusion, Dp (cm²/sec)-(mm Hg)	300°K	540	1015	520
	77°K	173	230	130
Recombination, α (cm³/sec)		$<10^{-10}$	1×10^{-8}	...
Ionization, β_i (cm³/sec)		...	...	1×10^{-9}
Conversion, K_c (cm⁶/sec)		6×10^{-32}	...	...

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

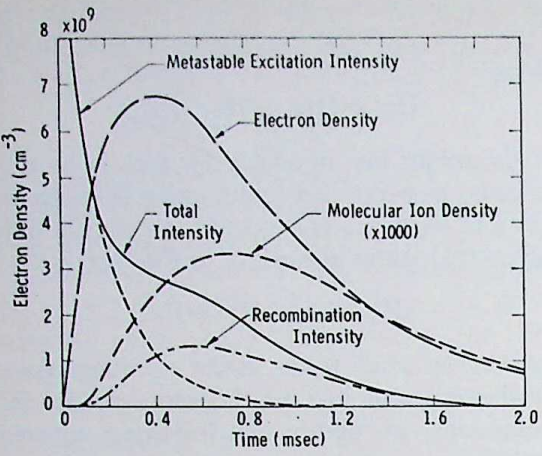


FIG. 4. Calculated behavior of various quantities in a helium afterglow at 300°K and a gas pressure of 0.54 mm Hg based on the model of the afterglow described in the text.

and atomic ions at the beginning of the afterglow, metastable-metastable ionizing collisions cause an increase in n_e and n_1. As the metastable density decreases, leading to a smaller ionization rate, the diffusion to the walls and conversion of atomic ions leads to the final decay of n_1 and n_e. Meanwhile, the molecular-ion density, which also started essentially from zero at the beginning of the afterglow, increases as a result of atomic ion conversion, reaching a considerably smaller maximum value than the electrons somewhat later in the afterglow.

One may account for the intensity variation of the afterglow radiation, as follows; in the early afterglow when metastable atoms are plentiful and molecular ions are absent the radiation should be the result of the production of fast electrons by metastable-metastable collisions which then excite other metastables to radiating states. Thus, the radiation produced by this process should decay at *three* times the metastable atom-decay rate.

Considering the electron's loss of kinetic energy at elastic collisions, it can be shown that approximately 6×10^3 collisions are required to reduce the electron's initial energy from ~15 eV to the threshold value, ~3 eV, for exciting $2\,^3S$ metastable atoms to the observed radiating states (e.g., $3\,^3D$). For the inferred initial metastable density of $\sim10^{12}$ cm^{-3}, the observed early radiation intensity can be accounted for, provided that the average cross section for excitation of the metastable state to a radiating state over the electron energy range from 3–15 eV is comparable to the cross section for scattering of electrons by normal atoms,[20] i.e., $\sim4\times10^{-16}$ cm^2. Although there are no direct measurements of electron impact excitation of metastable states, such a cross section does not seem unreasonable, since it is substantially less than the theoretical maximum value[21] for such a process. The predicted shape of the early afterglow radiation (metastable excitation intensity) is indicated in Fig. 4.

In the late afterglow, when the molecular ions have reached an appreciable concentration and the metastable atoms have diffused away, the radiation should be the result of excited states produced by dissociative recombination, reaction (13). The intensity of this radiation should, therefore, be roughly equal to $(\alpha n_e n_2)$, where α is the dissociative recombination coefficient, on the assumption that recombination predominantly produces the desired state either directly or by cascading from higher states. The recombination radiation intensity curve of Fig. 4 was predicted from the computed electron density and molecular-ion density curves and an assumed recombination coefficient of 1×10^{-8} cm^3/sec.[22] The total predicted intensity is then the sum of the metastable excitation and recombination intensities. The absolute intensities measured in the late afterglow are consistent with this model and the assumed recombination coefficient. However, the difficulties of the absolute intensity determinations are such that the values may be in error by up to a factor of 3.

It should be noted that, with the aforementioned model of the afterglow, the radiation resulting from metastable excitation by electron impact in the early afterglow should exhibit a thermal spectral linewidth. By contrast, the late afterglow radiation, which is assumed to result from the dissociative recombination process, should exhibit a broadened line shape.

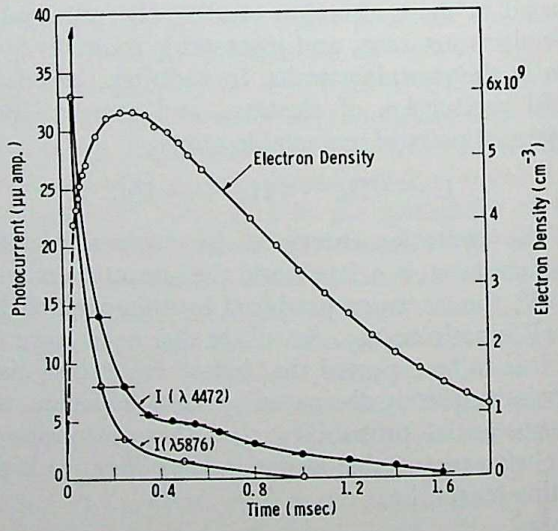


FIG. 5. Observed behavior of the electron density and the intensities of two lines emitted from a helium afterglow at 300°K and $p=0.54$ mm Hg. The line λ5876 corresponds to the $3\,^3D$–$2\,^3P$ transition; λ4472 to $4\,^3D$–$2\,^3P$.

[20] R. B. Brode, Rev. Mod. Phys. 5, 257 (1933).

[21] H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena* (Clarendon Press, Oxford, 1952), p. 144.

[22] This value was used as an approximation of the results of C. L. Chen, C. C. Leiby, and L. Goldstein, Phys. Rev. 121, 1391 (1961), who are the only investigators who achieved recombination control in their afterglow.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

V. RESULTS AND DISCUSSION

The observations which led to the construction of the afterglow model discussed in the previous section suggested a natural subdivision of the afterglow into an early and a late period, based on the line intensity and linewidth variations. In Figs. 5 and 6 are shown the variations of n_e and of the intensities of λ5876 and λ4472 observed during afterglows at $T=300°K$ and 0.54 and 2.05 mm Hg pressure, respectively. It will be seen that the observations at 0.54 mm Hg reproduce qualitatively the predictions of the model shown in Fig. 4, in that the electron density reaches a maximum value at ~300 μsec in the afterglow and the intensity of λ4472 exhibits a slight hump at ~500 μsec. No comparisons with the 2.05 mm Hg data (Fig. 6) are made, since the simplifying assumptions used in the model calculation at 0.54 mm Hg pressure are no longer valid at the higher pressure.

On the basis of the proposed afterglow model one should expect significant changes in the character of the radiation emitted in the early and in the late afterglow. It is noted that in the early afterglow the relative intensities of the various lines emitted are quite similar to those observed during the discharge. In both cases the excited states are presumed to be produced principally by electron impact excitation of lower excited states (ground state or metastable states). From the observed intensities in the early afterglow, one deduces a population of excited states which decreases monotonically with increasing principal quantum number, n. (Lines from upper states $n=3, 4, 5, 6$ were observed.) By contrast, in the late afterglow a change in spectral distribution is noted, the relative population of the states of higher quantum numbers, e.g., $n=5$, increasing markedly in comparison to the early afterglow values. As will be seen in the later discussion, such a behavior is consistent with dissociative recombination into higher excited states.

A principal effort of the present investigation lay in the determinations of the widths of the λ5876 line during the discharge and at various times in the afterglow in order to test the dissociative recombination hypothesis. As discussed in Sec. IIIC, the measured linewidths include contributions from fine structure, Doppler broadening, and instrumental broadening. The results of the line broadening measurements are, therefore, presented in terms of the observed total linewidths, determined at the half-maximum intensity points and expressed as fractions f of the free spectral range of the etalon, $\Delta\tilde{\nu}_{fsr}=(2\mu t)^{-1}=0.667$ cm^{-1}. The "temperature" of excited atoms required to yield the Doppler broadening contribution is then determined from the line-shape calculations as described in Sec. III C.

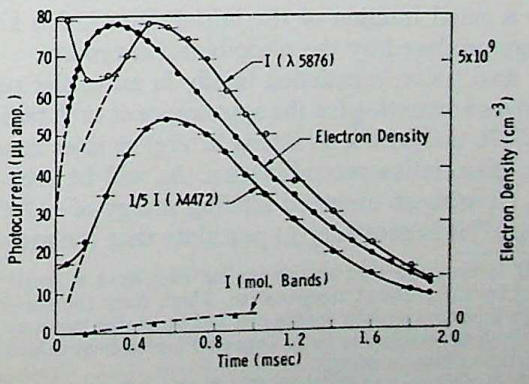


FIG. 6. Observed behavior of the electron density, two line intensities and the molecular band intensities in a helium afterglow at 300°K and $p=2.05$ mm Hg.

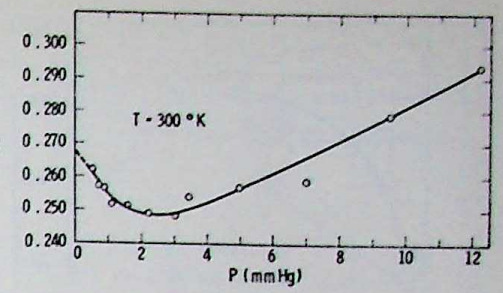


FIG. 7. Observed linewidth (including instrumental width) during the pulsed discharge in helium at 300°K. The linewidths are expressed as fractions f of the free spectral range of the Fabry–Perot etalon, $\Delta\tilde{\nu}_{fsr}=0.667$ cm^{-1}. The minimum width at ~3 mm Hg corresponds to the computed width for an excited atom temperature of ~300°K.

The observed width of the λ5876 line during the *discharge* in helium for various gas pressures is shown in Fig. 7. The minimum width observed at ~3 mm Hg corresponds to the linewidth calculated for an excited atom temperature of ~300°K. The increase in width at >3 mm Hg is the result of the appearance of some molecular light and pressure broadening effects. The increase in linewidth with decreasing pressure below 2 mm Hg can be shown to result from radiation of the excited atoms before they lose their recoil kinetic energy gained at the inelastic impact of the electrons having more than 23 eV kinetic energy which create the excited states. The linewidth data extrapolated to zero pressure are consistent with the expected recoil energy of ~0.004 eV, and the fact that this additional broadening does not disappear until ~3 mm Hg pressure suggests that excitation transfer to slow atoms and loss of kinetic energy by collisions of the excited atoms with normal atoms are slow enough to permit observation of line broadening due to dissociative recombination at pressures ≲2 mm Hg.[23]

The observed λ5876 linewidths in the afterglow, as well as during the discharge, are shown for several gas

[23] An alternative possibility is that narrowing of the discharge line may be due to a change in mode of production of excited states with change of gas pressure. If metastable $2\,^3S$ atoms reach a very large density $>10^{13}$ cm^{-3} as the pressure increases, then impact excitation to the $3\,^3D$ state by ~3 eV electrons will produce the observed intensity of λ5876 and the atom recoil energy will be too small to appreciably add to the thermal linewidth.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

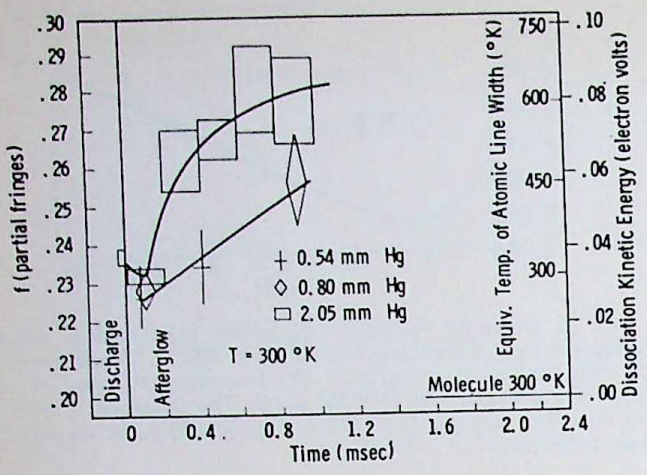


FIG. 8. Variation of the observed linewidths with time during the discharge and in the afterglow of helium at 300°K. The widths are expressed as fractions f of the free spectral range of the etalon. The scales at the right refer, respectively, to the equivalent atom temperature required to yield the observed width and to the dissociation kinetic energy E_D which must be added to the initial molecular ion thermal motion to give the observed width.

pressures at $T=300$°K in Fig. 8 and at $T=77$°K in Fig. 9. Here, improved Fabry-Perot plates were used. They had dielectric reflecting coatings of higher reflectivity and lower absorptivity than the silver films used on the previous plates. Thus, the instrumental broadening was reduced relative to Fig. 7. The horizontal widths of the symbols indicate the time gate during which the photomultiplier was made sensitive to light; the vertical extent of the symbols indicates the uncertainty in the linewidth determinations in these low-signal measurements.

From the temperature scale at the right of the figure, it will be seen that, as noted previously at 300°K, the line during the discharge at 77°K shows broadening in excess of ambient temperature effects, consistent with atom recoil at impact of the exciting electrons. Immediately afterwards, in the early afterglow at 300 and at 77°K, the linewidths decrease to values indistinguishable from the widths appropriate to the ambient temperatures. This observation is consistent with the proposed mechanism for production of the early afterglow radiation by electron impact on metastable atoms, since the ~3 eV electron energy required to excite to the radiating $3\,^3D$ state causes negligible recoil in the struck atom compared to its thermal speed.

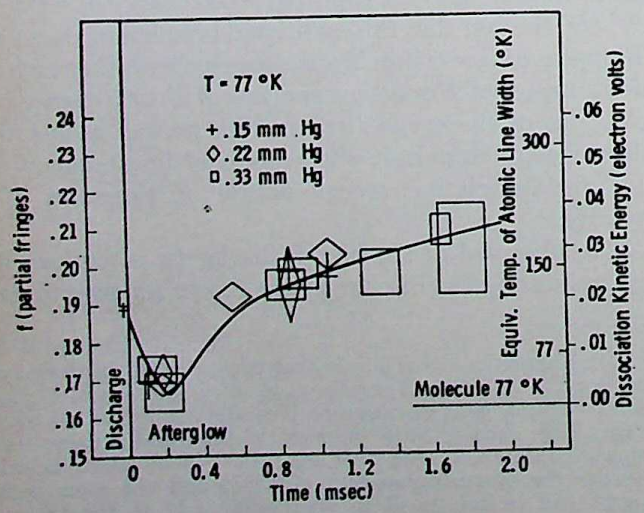


FIG. 9. Variation of the observed linewidths in helium at 77°K.

It will be seen in both Figs. 8 and 9 that the observed linewidths increase with time in the later afterglow. These observations are consistent with the expectation that, as the contribution to the radiation by metastable atom excitation decreases and that due to dissociative recombination increases, the line should increase in width, until a maximum width is attained when recombination alone is important. Unfortunately, the decrease in intensity with time in the afterglow does not permit line width measurements to be carried beyond ~1 msec at 300°K and ~2 msec at 77°K.

Although a pronounced line broadening is detected in the late afterglow, if one calculates the dissociation kinetic energy required to yield the observed width,[24] then, as the second scale at the right of the figures indicates, the inferred change of potential energy between the initial and final states of reaction (13) is rather smaller than one would estimate. As mentioned in Sec. III C.3, it is difficult to assure radiation of the excited, fast moving atom before slowing down by elastic collisions or excitation transfer to slow atoms takes place. If one argues that the momentum transfer cross section for scattering of $3\,^3D$ atoms by normal helium atoms is the same magnitude as for the metastable $2\,^3S$ atoms, namely, $Q_m=8\times10^{-15}$ cm^2,[18,19] then the excess energy of the atoms produced by dissociative recombination decays at the rate $\nu_{coll}\sim10^8$ sec^{-1} at 2 mm Hg pressure. Alternatively, if the excitation transfer cross section is of the same order of magnitude ($\sim10^{-14}$ cm^2) as the momentum transfer cross section, appreciable "narrowing" of the observed line will also occur by the transfer process. Excitation transfer cross sections, even for nonresonant cases, attain large values. For example, Gabriel and Heddle[25] deduce values of $Q_{transfer}\gg10^{-15}$ cm^2 for various $n\,^1P\rightarrow n\,^1D$ transfers ($n=3, 4, 5, 6$) at thermal energy. Thus, it is quite possible that in these studies we have been able to observe only a small fraction of the initial dissociation kinetic energy produced by the recombination process.

A final point, mentioned briefly in an earlier paper,[3] involves accounting for the appearance of such radiation as λ5876 and λ4472 as strong afterglow lines resulting from dissociative recombination. As will be seen from Fig. 10, with an expected binding energy of ~2 eV for (He_2^+),[26] it is necessary to postulate that the molecular

[24] It is assumed that the molecular ions have a speed determined by the ambient temperature. Then, since the dissociation energy is shared equally by the two helium atoms, one may calculate the line broadening to be expected for various values of the dissociation kinetic energy.

[25] A. H. Gabriel and D. W. O. Heddle, Proc. Roy. Soc. (London) A258, 124 (1960).

[26] E. A. Mason and J. T. Vanderslice, J. Chem. Phys. 29, 361 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ions are in vibrationally excited states as they undergo dissociative recombination in order to produce the high-lying atomic states which yield the observed lines. (The figure is schematic, the vibration levels having been calculated for a Morse potential function.) The formation of the molecular ions in high-lying vibration states is a reasonable consequence of the assumed conversion reaction (12), in which the most likely binding collision is one in which a minimum loss of kinetic energy between two atoms takes place. Recent studies of vibration persistence in gases[27] have indicated that, for such well-separated vibration levels as in He_2^+, the probability of de-excitation by transfer of vibration energy to kinetic energy of relative motion at a collision with a helium atom is likely to be very small, $<10^{-8}$. Thus, on the millisecond time scale of our measurements, the ions should persist in the higher vibration states for an appreciable fraction of the observing interval.

In summary, the initial goal of determining whether or not the dissociative process is responsible for the large electron-ion recombination losses observed in microwave afterglows has not been fully achieved. The observations are consistent with a model of the helium afterglow in which, at later times, recombination becomes an important source of the afterglow radiation. At these times the emitted lines definitely show broadening in excess of thermal width. If one were to argue that the late afterglow broadening is the result of some process other than dissociative recombination, then one is faced with the dilemma of explaining how the excited atom temperature in the early afterglow sinks to thermal values and then increases rather abruptly. Clearly, such effects as pressure broadening cannot explain these observations, and careful monitoring of molecular band

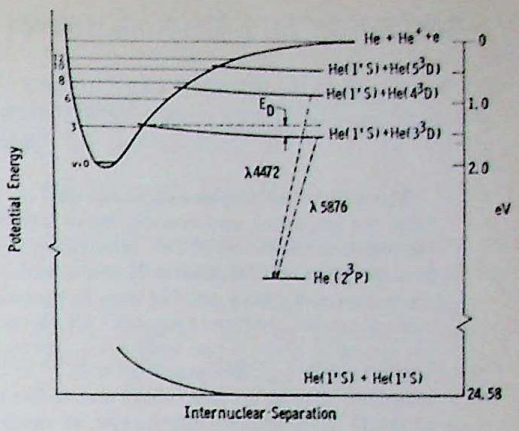


FIG. 10. Schematic representation of the potential curves required to explain the observations of strong λ4472 and λ5876 afterglow lines in terms of dissociative recombination. It will be seen that long-lived vibrationally excited states of the molecular ion are involved.

radiation intensities indicates that such spurious broadening effects do not account for the observed increase in width. Thus, dissociative recombination appears the most likely source of the modestly increased linewidths in the late afterglow. Studies of line shapes in neon afterglows, where recombination is the predominant electron removal process over a wider range of experimental variables, are in progress.[28]

ACKNOWLEDGMENTS

The authors wish to thank the members of the Atomic Physics Group, especially L. S. Frost and A. V. Phelps, for their interest and helpful discussions of this problem.

[27] S. J. Lukasik and J. E. Young, J. Chem. Phys. 27, 1149 (1957).

[28] T. R. Connor and M. A. Biondi, Sixteenth Gaseous Electronics Conference, Pittsburgh, Pa., October, 1963 (unpublished); Bull. Am. Phys. Soc. 9, 184 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Ionization and Dissociation of Fast H_2 Molecules Incident on H_2 Gas*

G. W. McClure
Sandia Laboratory, Alburquerque, New Mexico
(Received 9 January 1964)

The yield cross sections $\sigma_{H_2^+}$, σ_{H^+}, and σ_H for production of fast H_2^+ and H^+ ions and H atoms resulting from the ionization and dissociation of fast H_2 molecules in single collisions with H_2 gas molecules have been measured in the 6- to 120-keV energy range (collision velocity 0.7 to 3.2×10^8 cm/sec). The neutral primary beam contained a fraction of H atoms which was determined by proportional counter analysis. Secondary particles produced by the H atoms in the primary beam were determined in separate H beam experiments and subtracted from the composite (H,H_2) neutral beam data to yield the production of secondaries due to H_2 primaries alone. Cross section $\sigma_{H_2^+}$ increases monatonically from 0.5 to 2.0 Å^2 in the energy range 6 to 120 keV while σ_{H^+} increases from 0.28 Å^2 at 7.5 keV to a maximum of 0.70 Å^2 at 50 keV. Cross section σ_H, measured only at 10 keV, was found to have an approximate value of 2.5 Å^2 based on an assumed shape of the H atom angular distribution at small angles. The results are compared with data on a number of collision processes in the same velocity range involving various species of hydrogen ions and H atoms incident on H_2 gas molecules.

INTRODUCTION

THERE is a growing need for experimental data concerning ionization and charge exchange in collisions of atomic particles at relative velocities of the order of 10^8 cm/sec. Such velocities are encountered in fusion machines, ionic propulsion engines, high-voltage gaseous dischanges, and nuclear explosions. Experiments with hydrogen ions and molecules are directly applicable to all of these problems and are needed as a guide to further development of collision theory. In the particular velocity range of interest, theoretical calculations of collision cross sections are most difficult because the optical electron velocity of the colliding particles is about the same as the atomic translational velocity. Neither the adiabatic nor the impulsive model of a collision is valid under these circumstances.

The main purpose of this paper is to report measurements of the cross sections for ionization and dissociation of 6- to 120-keV H_2 molecules in collision with H_2 gas molecules. Apparatus used previously for the study of H^+, H_2^+, and H_3^+ collisions with H_2 molecules was employed.[1] The method consists of preparing a monoenergetic beam of fast molecules by charge exchange, passing the beam through a thin target of H_2 gas and measuring the production of fast H^+ and H_2^+ ions and H neutral atoms.

The reactions contributing to the production of the observed species are as follows, neglecting reactions in which negative ions are formed.

$$\text{Reaction 1:}\quad \overset{\circ}{H_2}+H_2\rightarrow H+H+(H_2),$$
$$\text{Reaction 2:}\quad H_2+H_2\rightarrow H_2^+ +(e+H_2),$$
$$\text{Reaction 3:}\quad H_2+H_2\rightarrow H^+ +H+(e+H_2),$$
$$\text{Reaction 4:}\quad H_2+H_2\rightarrow H^+ +H^+ +(e+e+H_2).$$

Here the heavy fragments of the primary H_2 molecules are shown immediately following the arrow while electrons removed from the primary and the post-collision target molecule are shown in parentheses on the right. The target H_2 molecules may be ionized, excited or dissociated, and the product particles H_2^+ and H may carry off some energy in the form of internal excitation, but in the great majority of collisions of the type considered the two constituent protons of the primary molecules, whether bound or dissociated after the collision, suffer very little change of velocity or direction of motion.

The present measurements do not completely resolve the above reactions but rather determine the yield cross sections σ_{H^+}, $\sigma_{H_2^+}$, and σ_H for the three types of fast heavy particles H^+, H_2^+, and H. In terms of the reaction cross sections σ_1, σ_2, σ_3, σ_4 for the above reactions, the measured quantities are:

$$\sigma_{H_2^+}=\sigma_2,$$
$$\sigma_{H^+}=\sigma_3+2\sigma_4,$$
$$\sigma_H=2\sigma_1+\sigma_3.$$

Particle counting was used to measure the intensities of both the primary neutral beams and the secondary ion and neutral beams. The counter employed permitted H_2 and H neutrals of the same velocity to be clearly distinguished by the relative size of electrical impulses they produced. Ability of the counting system to resolve H and H_2 counts placed a lower limit of about 3 keV on the present method of measurement. The narrow slit aperture on the proportional counter (needed to support a thin-film gas barrier) prevented our use of a technique devised by Sweetman[2] which might otherwise have been employed to distinguish the reactions 1 to 4.

The apparatus was designed so that all heavy fragments of the primary molecule having a post-collision angle less than 2.2 deg relative to the direction of the H_2 beam could be detected. At the 6-keV lower energy limit of the present (H_2+H_2) measurements the cross section for Coulomb scattering through an angle greater

* This work performed under the auspices of the U. S. Atomic Energy Commission.
[1] G. W. McClure, Phys. Rev. 130, 1852 (1963).
[2] D. R. Sweetman, Proc. Roy. Soc. (London) A256, 416 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

than 2.2 deg of one of the two protons of the incident molecule against one of the two protons of the target molecule is 2×10^{-17} cm^2, neglecting screening. This cross section is of the same order of magnitude as the measured H^+ production cross section at 6 keV; hence, some error may be introduced into the H^+ data at this energy. However, throughout most of the present energy range no significant large angle scattering errors are expected due to Coulomb interactions.

At angles up to a few degrees from the forward direction the differential cross section for production of dissociation fragments of fast molecules may exceed that due to Coulomb scattering because of electronic transitions of the incident particle to antibonding states in which the two nuclei repel each other. Resulting production of dissociation fragments at angles outside the detector aperture may cause the measured cross sections to be 10 to 20% below the true total cross section at the lowest energies.

An important problem associated with the present measurements was that of producing a primary beam of neutral H_2 molecules. This was accomplished via the H_2^+ on H_2 reactions, studied previously,[1,2] which give rise to a mixed beam of H_2 molecules and H atoms of the same velocity via the processes

$$\text{Reaction 5:}\quad H_2^+ + H_2 \rightarrow H_2 + (H_2^+),$$

$$\text{Reaction 6:}\quad H_2^+ + H_2 \rightarrow H + H + (H_2^+),$$

$$\text{Reaction 7:}\quad H_2^+ + H_2 \rightarrow H + H^+ + (H_2).$$

In order to calculate the background production of H^+ caused by electron loss from the H atoms in the mixed neutral beam, it was necessary to know the cross section $\sigma_{0,1}$ for the reaction

$$\text{Reaction 8:}\quad H + H_2 \rightarrow H^+ + (e + H_2).$$

Although this cross section has been measured,[3–9] a serious discrepancy existed between the available measurements below 10 keV. Therefore, $\sigma_{0,1}$ was remeasured over the entire energy range 2 to 120 keV. At the same time, remeasurements of $\sigma_{0,-1}$, the cross section for the reaction

$$\text{Reaction 9:}\quad H + H_2 \rightarrow H^- + (H_2^+)$$

were made to resolve a gross discrepancy in previous data on this process. Both measurements involving H atom primaries are discussed in the first portion of the section entitled Results. The second portion of the Results section deals with the (H_2+H_2) measurements.

All cross sections were measured relative to the cross section $\sigma_{1,0}$ for the reaction

$$\text{Reaction 10:}\quad H^+ + H_2 \rightarrow H + (H_2^+).$$

This reaction[10] serves as an excellent basis for calibration of collision experiments of the present type because (1) no excited states are possible in the incident proton beam, (2) the angular distribution of the product H atoms is very strongly peaked in the forward direction, and (3) the absolute value of the cross section has been measured in the neighborhood of 10–100 keV by a number of investigators whose results are in excellent agreement.

A major problem in the interpretation of any experiment using fast neutral beams produced by charge exchange is that of uncertainty concerning the quantum states of the incident particles. In the present work this question is not settled; however, a discussion of the possible states of excitation of the neutral particles is given. Also, the conditions of production of the neutral beams are fairly well defined so that future data on the initial population and decay of excited states produced in charge-exchange collisions may be brought to bear on a more refined interpretation of the present work.

The section entitled Discussion compares the (H_2+H_2) collision data, obtained for the first time in this investigation, with data on related collision processes.

APPARATUS

The ion source, ion accelerator, and magnetic beam analyzer described in Ref. 1 were used without alteration to provide a 1–20-keV beam of either H^+ or H_2^+ ions having an energy spread of approximately one percent. Either of these beams could be directed along the sequence of apertures S_1, S_2, S_3, and S_4 (Fig. 1) which formed the entrance and exit slits of two differentially pumped chambers T_1 and T_2 operated with separate gas controls and pressure gages. These chambers served as a neutralizing chamber and main collision chamber, respectively. Electrostatic deflection plates D_1 located between T_1 and T_2 were used in the neutral beam experiments to deflect unconverted ions emergent from T_1 out of the neutral beam, permitting only neutrals to enter T_2. To change from an ion beam experiment to a neutral beam experiment required only the closing of a bypass valve between T_1 and the main vacuum chamber, the admission of a gas to the neutralizing chamber T_1, and the application of a deflection voltage to plates D_1.

A second set of electrostatic deflection plates D_2 at the

[3] J. H. Montague, Phys. Rev. **81**, 1026 (1951).
[4] P. M. Stier and C. F. Barnett, Phys. Rev. **103**, 896 (1956).
[5] C. F. Barnett and H. K. Reynolds, Phys. Rev. **109**, 355 (1958).
[6] I. M. Fogel, V. A. Ankudinov, D. V. Pilipenko, and N. V. Topolia, Zh. Eksperim. i Teor. Fiz. **34**, 579 (1958) [English transl.: Soviet Phys.—JETP **7**, 400 (1958)].
[7] D. V. Pilipenko and I. M. Fogel, Zh. Eksperim. i Teor. Fiz. **42**, 936 (1962) [English transl.: Soviet Phys.—JETP **15**, 646 (1962)]. Data on $\sigma_{0,1}$ from this reference (9–30 keV) agree with data from Ref. 6.
[8] R. Curran and T. M. Donahue, Phys. Rev. **118**, 1233 (1960).
[9] Fred Schwirtzke, Z. Physik **157**, 510 (1960).
[10] S. K. Allison and M. Garcia Munoz, in *Atomic and Molecular Processes*, edited by D. R. Bates (Academic Press Inc., New York, 1962), p. 751.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

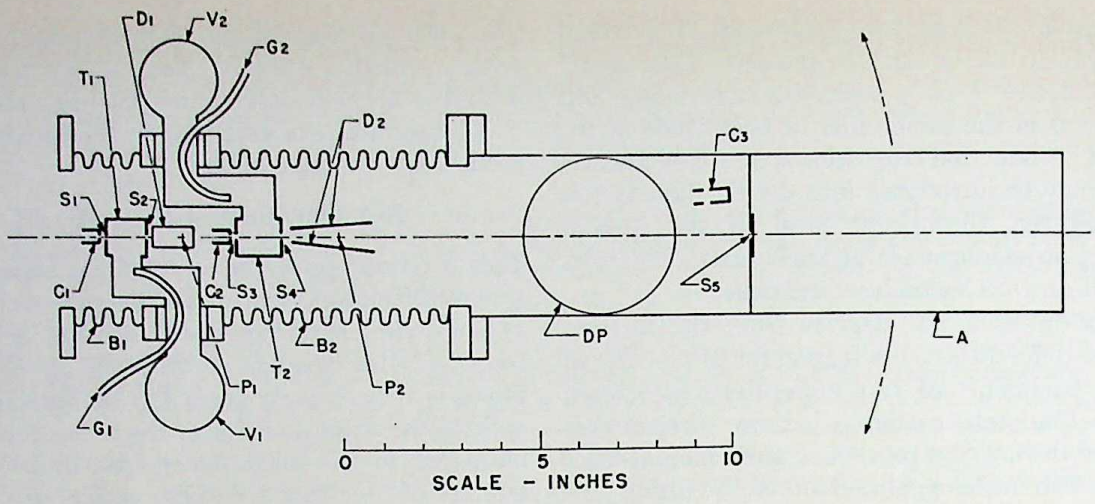


FIG. 1. Apparatus diagram. T_1 and T_2, collision chambers. V_1 and V_2, ionization gauges. G_1 and G_2, gas inlets. C_1, C_2, and C_3, Faraday cups. D_1 and D_2, deflection plates for deflecting charged particle beams off the main beam axis. S_1, S_2, S_3, and S_4, circular apertures having diameters 0.01, 0.02, 0.01, and 0.10 in., respectively. S_5, rectangular detector slit 0.001 in.×0.963 in.; long dimension perpendicular to plane of drawing. A, ionization chamber forming lower section of counter assembly shown in Ref. 1. DP, 400 liter/sec oil diffusion pump. B_1 and B_2, bellows-type vacuum joints constrained by external gimbals. B_1 is used to align apertures S_1 and S_3 with direction of primary beam incident on S_1 from the left. B_2 is used for mechanically sweeping detector slit S_5 across the electrostatically analyzed set of particle beams emergent to the right from D_2. B_2 permits detector slit S_5 (and attached assemblies A and DP) to swing through a $\pm 20°$ arc in the plane of the drawing about center of rotation P_2.

exit of T_2 divided the different species of fast charged particles emergent from T_2 into separate beams. Each of these beams was measured by moving the entrance slit S_5 of the proportional counter detector across the beam while recording with single-channel pulse-height analyzers the integrated counts of "single" and "double" amplitude pulses. The single-amplitude pulses came from H^+ or H particles and the double-amplitude pulses came from H_2 or H_2^+. Counts occurred in both channels only when the neutral beam from primary H_2 neutrals was being scanned. In this case pulse-height discrimination was used to distinguish H from H_2 in the composite neutral beam emergent from the collision chamber. The geometry and properties of the proportional counter were discussed thoroughly in Ref. 1. A 256-channel pulse-height analyzer was used in setting the pulse-height acceptance limits of the single-channel analyzers.

A refinement of the original apparatus[1] was the addition of a servo mechanism to drive the counter entrance slit at a speed proportional to the primary ion beam current. This system greatly reduced errors due to beam current fluctuations. The servo drive signal was derived in the neutral beam measurements from a Faraday cup C_2 surrounding the entrance of chamber T_2 into which the unconverted ion beam emergent from T_1 was directed. In all cases only a small fraction of the beam entering T_1 was neutralized and all unconverted ions were received by the Faraday cup. High stability of the neutralizing chamber pressure was obtained with a specially designed gas leak.[11] Constancy of this pressure insured that the neutral beam entering T_1 was exactly proportional to the servo drive signal once the leak was set.

[11] D. L. Allensworth, Rev. Sci. Instr. 34, 448 (1963).

CROSS-SECTION DETERMINATION

Cross sections were determined from the formula

$$\sigma_i(j)=I_i/I_j nd, \tag{1}$$

where $\sigma_i(j)$ is the cross section for production of secondary particles of type i by primaries of type j, I_j is the current of primary particles of type j, I_i is the current of secondary particles of type i, and nd is the equivalent number of gas molecules in the collision chamber per cm^2 of target area normal to the beam direction.

Because molecular flow conditions were present in the collision chamber the following relation holds

$$nd=\alpha p, \tag{2}$$

where p is the collision chamber pressure gage reading and α is a calibration constant having a fixed value for a given target gas and a given statistical distribution of incident-particle paths through the collision chamber. It is assumed that the distribution of ray paths is independent of the type and energy of primary beam used in these experiments.

With the further assumption that

$$N_i/N_j=I_i/I_j, \tag{3}$$

where N_i is the number of counts recorded while the detector is scanned across the type i secondary beam and N_j is the number of counts recorded while the detector is scanned across the type j primary particle beam, Eqs. (1)–(3) yield the following cross-section formula:

$$\sigma_i(j)=(N_i/N_j)(1/\alpha p). \tag{4}$$

Secondary particles produced at the edges of the entrance slit to chamber T_2 or "spurious secondaries"

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which are present in the incident beam entering the collision chamber are not dependent on the collision chamber pressure p. Either of these effects or the presence in T_2 of a residual gas or vapor having a different cross section than the target gas result in failure of Eq. (4) to give a correct cross section. All three sources of error can be eliminated by plotting N_i versus p and by using the slope $\Delta N_i/\Delta p$ of this curve in place of N_i/p in evaluating Eq. (4). This procedure was used in every cross-section determination of the present work.

To summarize, the general working cross-section formula is

$$\sigma_i(j)=(1/\alpha N_j)(\Delta N_i/\Delta p). \quad (5)$$

CALIBRATION

The calibration of the apparatus consisted of determining a single value of the gas-dependent constant α of Eq. (5) since all measurements were made with H_2 as the target gas. This constant was determined by measuring the conversion of protons to H atoms by electron capture in H_2 gas (reaction 10) and assuming the cross-section value 8.2×10^{-16} cm^2 at the calibration proton energy of 10 keV. The value of α was found to be constant within a range of $\pm5\%$ in several determinations made during the course of the work. The chief cause of variations in successive α determinations was beam current fluctuations. The servomechanism used to compensate for beam current variations in all of the other measurements herein reported could not be used for the calibration because when the H^+ beam current was made low enough that the primary beam could be counted directly at the output of T_2, the fringing portion of the beam entering Faraday cup C_2 was too weak to provide a monitoring signal to drive the servo properly. This difficulty was offset in the calibration by taking a large number of points on the N_i versus p curves.

The use of reaction 10 for calibration continues the practice begun during our earlier work of reducing all hydrogen target measurements to a common basis. Should the value assumed for this cross section prove to be in error in the future, all of our measurements can be corrected accordingly by applying the same multiplicative correction factor throughout.

EXPERIMENTAL CONDITIONS

Hydrogen gas at a pressure between 10^{-4} and 10^{-3} mm Hg was used as the neutralizing gas in all of the neutral beam measurements. These pressures were sufficiently low that less than 10% of the incident ions were neutralized. The residual gas pressure in the neutralizing chamber was about 10^{-5} mm Hg with the gas inlet valve closed and about 10^{-6} mm Hg with the bypass valve open for calibration measurements.

Hydrogen gas of 99.9% purity was used in the collision chamber as the target gas. The residual gas pressure in the main collision chamber was 5×10^{-7} mm Hg with the hydrogen inlet leak closed and was varied between 10^{-5} and 10^{-4} mm Hg in order to obtain

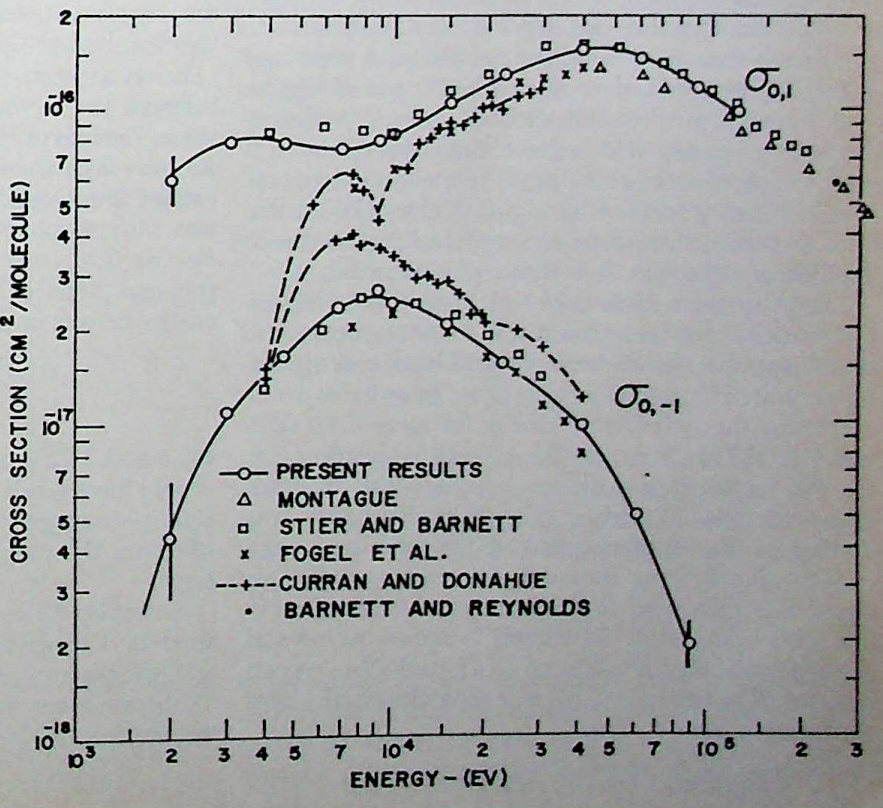


FIG. 2. Cross section $\sigma_{0,1}$ for conversion of primary H atoms to H^+ ions and cross section $\sigma_{0,-1}$ for conversion of primary H atoms to H^- ions. Present results are compared with previous data of Montague (Ref. 3), Stier and Barnett (Ref. 4), Barnett and Reynolds (Ref. 5), Fogel *et al.* (Refs. 6 and 7), and Curran and Donahue (Ref. 8).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the N_i versus p curves. The maximum fraction of incident particles which suffered either charge changing of dissociative collisions in the collision chamber was 0.7% and the maximum fraction of either the primary or secondary particles which underwent charge changing or dissociative collisions in the drift space between the collision chamber and the detector at the maximum drift space pressure of 10^{-6} mm Hg was about 0.07%. Under these conditions the attenuation of both the primary beam and secondary beams during traversal of the collision chamber and drift space could be ignored in evaluating the quantities N_i and N_j. Direct evidence of thin-target conditions was seen in the fact that all of the N_i versus p curves were linear within the statistical uncertainties of the particle counts. This uncertainty varied from 2 to 10%.

Ion beam currents entering chamber T_1 were generally of the order of 10^{-13} A throughout the data runs.

RESULTS

H Primaries

By permitting reaction 10 to occur in chamber T_1 and by applying a suitable deflection voltage to plate D_1 a beam of H atoms cleared of H^+ and H^- ions was directed into the collision chamber T_2. The fast secondaries H^+ and H^- emergent from T_2 due to reactions 8 and 9 were deflected by plates D_2 at angles of ± 3 deg from the unconverted H primary beam. Scans of the detector across all three beams at several energies indicated that a 1-deg scan was sufficient to receive essentially all of the particles emergent from the 4.4-deg-wide collision chamber exit aperture S_4. Consequently, 1-deg scans were used in all the $\sigma_{0,1}$ and $\sigma_{0,-1}$ measurements. For determinations of $\sigma_{0,1}$ the H^+ and H beams were scanned by the detector at several collision chamber pressures. Throughout these scans the H^+ beam emergent from chamber T_1 was collected at cup 2 and used as a servo drive signal. Calculation of the cross section $\sigma_{0,1}$ was then accomplished by the direct application of Eq. (5). Results are plotted in Fig. 2.

Except in the neighborhood of 10 keV the yield of H^- particles from reaction 9 was found to be so low that a large increase in the primary H beam was needed for accurate H^- yield determinations. In order to avoid correcting the recorded H counts for large dead-time losses at the required increased beam intensities, cup C_3 was moved into position to collect the H beam emergent from T_2 and a suitable positive bias was applied to the suppressor ring at the entrance to the cup to give a large secondary electron signal. This signal was then used to drive the servo while the H^+ and H^- beam were alternately scanned at several pressures to obtain plots of N_{H^+} and N_{H^-} versus pressure. The ratio $\sigma_{0,-1}/\sigma_{0,1}$ was then determined from the formula

$$\sigma_{0,-1}/\sigma_{0,1}=[\Delta N_{H^-}/\Delta p]/[\Delta N_{H^+}/\Delta p] \qquad (6)$$

and $\sigma_{0,-1}$ was determined by multiplying this ratio by the values of $\sigma_{0,1}$ from Fig. 2. The results for $\sigma_{0,-1}$ are shown also in Fig. 2.

Uncertainties in the experimental points are $\pm 5\%$ except for the points at 2 and 90 keV where larger uncertainties are indicated.

Previous measurements of $\sigma_{0,1}$ and $\sigma_{0,-1}$ are shown in Fig. 2 together with the present results. At all energies from 4 to 120 keV except in the vicinity of 6 keV our results appear to be consistent with the results of Stier and Barnett.[4] Agreement is excellent for both $\sigma_{0,1}$ and $\sigma_{0,-1}$.

In the energy range 40 to 120 keV our data points and those of Stier and Barnett for $\sigma_{0,1}$ fall about 15% above those of Montague.[3] Between 15 and 40 keV the results of all investigators scatter over a range of $\pm 20\%$ from the mean, but smooth curves drawn through the individual sets of data possess very nearly the same shape. It seems likely, therefore, that the discrepancies in this region could be accounted for by uncertainties in collision chamber pressure determinations.

Schwirtzke[9] has determined a quantity he calls $\sigma_{0,1}$ which we believe is actually $\sigma_{0,1}+\sigma_{0,-1}$. The Schwirtzke result, thus interpreted, has been compared with values of $(\sigma_{0,1}+\sigma_{0,-1})$ calculated from our data and is found to agree within $\pm 15\%$ over the range 10–50 keV.

Below 10 keV the experimental values of $\sigma_{0,1}$ fall into two divergent groups with relatively high values given by the present results and those of Stier and Barnett, and relatively low values given by Curran and Donahue[8] and by Fogel *et al.*[6,7] The Fogel measurements do not extend below 7 keV where an extremely serious departure sets in between the present "high" values and the "low" values of Curran and Donahue.

In an attempt to understand the drastic departure between the $\sigma_{0,1}$ results of Curran and Donahue and the $\sigma_{0,1}$ values of the present investigation below 7 keV we have hypothesized the presence of the well-known excited $2s$ metastable state in our H atom beam which was not completely quenched by the relatively low electric field used between deflection plates D_1. This proposal seemed worthy of investigation since the energy defects in the two reactions

$$H(1s)+H_2 \rightarrow H^+ + (H_2+e),$$
$$H(2s)+H_2 \rightarrow H^+ + (H_2+e)$$

(13.6 and 3.4 eV, respectively) are such that the latter should have a relatively high cross section. Thus even if the percentage of atoms in the $2s$ state were small, the effect on the weighted average value of $\sigma_{0,1}$ could be large.

According to a private communication,[12] the fields used by Curran and Donahue[8] were large enough to give complete quenching of the $2s$ metastable; however, deliberate steps were not taken in our measurements to suppress the $2s$ states completely. The fraction of $2s$

[12] T. M. Donahue (verbal communication, 1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

atoms surviving after passing through the deflection field at D_1 was calculated from the expression

$$\phi = \exp\left[-\int_0^b \nu(x)\,(dx/v)\right], \qquad (7)$$

where ϕ, the survival probability, is defined as the ratio of the $2s$ atom current leaving the D_1 deflection field region to that entering the D_1 deflection region from chamber T_1, $\nu(x)$ is the local decay rate of $2s$ atoms at distance x from slit S_2, v is the H atom velocity, and b is the length of path from S_2 to the end of the deflection plates. The quantity $\nu(x)$ was calculated from the formula $\nu(x) = A[E(x)]^2$ where E is the estimated electric field strength at distance x and A is a constant equal to 2780 $\text{sec}^{-1}\ \text{V}^{-2}\ \text{cm}^2$.[13] [Equation (7) is based on the assumption that the incremental fraction of $2s$ atoms decaying in time increment dt is equal to the product νdt where ν is an instantaneous decay rate given by the "static-field" formula quoted.]

For the set of points plotted in Fig. 2 the calculated $2s$ survival probabilities are given in Table I. The rather irregular variation of the survival probability is caused by the extreme sensitivity of this quantity to the deflection voltage coupled with the fact that no special attempt was made to choose deflection voltages which would lead to a regular variation of ϕ. In some cases, particularly at low energies, the survival probabilities were somewhat large; hence a subsidiary experiment was performed as a specific test of a possible $2s$ effect.

Two measurements of $\sigma_{0,1}$ were made at 4 keV (not plotted in Fig. 2) in which the deflection voltages used on plates D_1 were 85 and 25 V, respectively, and for which the survival probabilities were 0.02 and 0.71, respectively. The two results were $\sigma_{0,1}(85\ \text{V}) = 7.92 \pm 0.4$ and $\sigma_{0,1}(25\ \text{V}) = 8.28 \pm 0.5$, indicating that the maximum increase in the cross section $\sigma_{0,1}$ due to $2s$ atoms would be about 20%. Therefore, a variable population of the $2s$ level does not seem to afford an explanation for the discrepancy with the Curran-Donahue results.

TABLE I. Calculated survival probabilities [ϕ, Eq. (7)] of $2s$ metastable atoms traversing deflection plates D_1 in the $\sigma_{0,1}$ measurements as a function of H atom energy.

Energy keV	Survival probability
2	0.145
3	0.206
4.5	0.276
6.8	0.095
9	0.270
10	0.575
15	0.056
22.5	0.10
40	0.0057
66	0.4×10^{-4}
120	10^{-3}

[13] H. Bethe and E. Salpeter, *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, Germany, 1957), Vol. XXXV, p. 373.

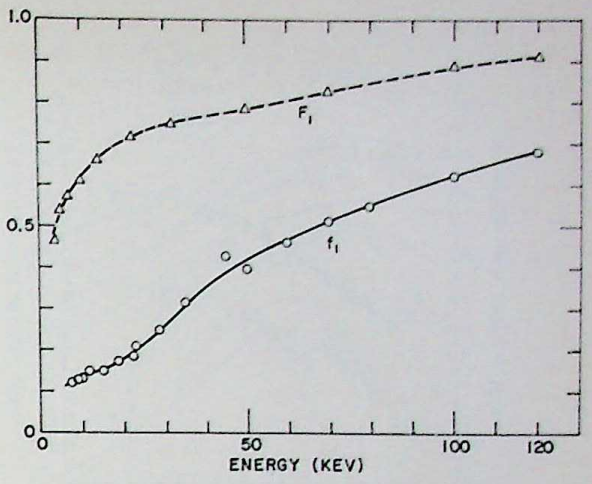


FIG. 3. Ratio $f_1 = N_H/(N_H + N_{H_2})$, where N_H and N_{H_2} are, respectively, the fluxes of H atoms and H_2 molecules in the composite neutral beam entering collision chamber T_2 during the H_2 primary measurements. Ratio $F_1 = \sigma_H/(\sigma_H + \sigma_{H_2})$, where σ_H and σ_{H_2} are, respectively, the cross sections for production of fast H and H_2 secondaries in collisions of fast H_2^+ ions with H_2 molecules. F_1 is calculated from the experimental data of Ref. 1.

H_2 Primaries

A composite neutral beam comprising fast H_2 molecules and H atoms was formed by passing magnetically analyzed H_2^+ ions into the neutralizing chamber. Fast H_2 and H neutrals produced in reactions 5, 6, and 7 emerged from the neutralizing chamber and entered the collision chamber while the emergent H_2^+ and H^+ ions were deflected out of the beam by deflector D_1.

The intensity and composition of the neutral beam emergent from T_2 was determined by making detector scans 1 deg wide. These scans were shown to be ample to cover the entire beam by trial runs at various widths. (The maximum angle of divergence of a primary neutral from the collision chamber axis was 0.38 deg as determined by the sizes of apertures S_2 and S_3.) Figure 3 shows the composition of the composite neutral beam emergent from T_2 as a function of the H_2 beam energy. The composition is described by the quantity

$$f_1 = N_H/[N_{H_2} + N_H], \qquad (8)$$

where N_H and N_{H_2} are, respectively, the H and H_2 counts recorded in a scan of the neutral beam. It may be seen that the composition varies considerably with beam energy and follows a form quite different from the function F_1 given by

$$F_1 = \sigma_H/[\sigma_H + \sigma_{H_2}], \qquad (9)$$

where σ_H and σ_{H_2} are, respectively, the production cross sections for H and H_2 secondaries in the reactions 5–7 calculated from the data of Ref. 1. This difference between f_1 and F_1 results from the small solid angle

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

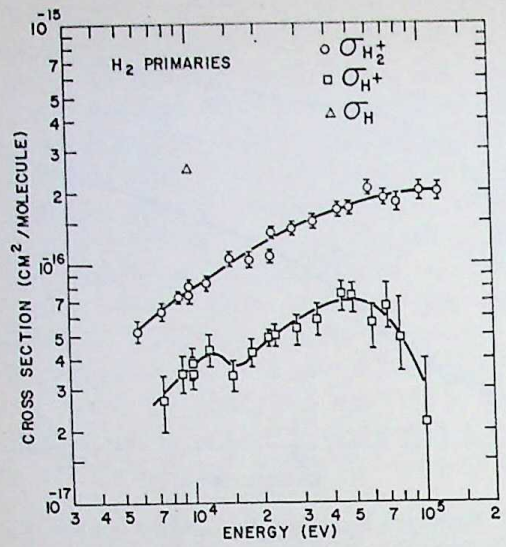


FIG. 4. Cross sections for conversion of fast H_2 molecules into fast secondaries H^+, H_2^+, and H in collisions with H_2 molecules. σ_{H^+}, $\sigma_{H_2^+}$, and σ_H are, respectively, the production cross sections of the secondaries H^+, H_2^+, and H.

subtended by aperture S_3 at the neutralizing chamber coupled with the fact that the H_2 molecules have a more strongly forward-peaked differential angular distribution than do the H dissociation products.

Because of the very narrow angular spread of the H_2 beam and the small size of apertures S_1 and S_3, it was necessary to carefully align the S_1–S_3 axis in order to achieve a minimum value of f_1 needed for optimum H_2 measurement accuracy. Owing to unexplained slight differences in the direction of the H_2^+ beam entering S_1 for various energies, care was taken to achieve good alignment at every accelerator energy. The slight departure of the point at 45 keV from the smooth curve f_1 probably resulted from misalignment which slightly favored H atoms at this energy.

To determine the cross sections for H_2^+ and H^+ production from H_2 primaries, a deflection voltage was applied to plates D_2 which deflected the secondary H_2^+ beam 3 deg from the main axis and the H^+ beam 6 deg. The neutral beam and the H_2^+ and H^+ beams were scanned repeatedly at several collision chamber pressures, providing curves of $N_{H_2^+}$ versus p and N_{H^+} versus p and values of N_H and N_{H_2}. Scan widths were adjusted to achieve complete collection of the H^+ and H_2^+ secondaries at each energy. One-deg scans sufficed for H_2^+ at all energies and 2-deg scans were sufficient for H^+ at all but the lowest energies. Below 10 keV the full 4.4-deg scans were used and considered to be adequate. The cross section $\sigma_{H_2^+}$ for H_2^+ production was straightforwardly determined from the formula

$$\sigma_{H_2^+}=-\frac{1}{\alpha}\frac{1}{N_{H_2}}\frac{\Delta N_{H_2^+}}{\Delta p} \qquad (10)$$

obtained from Eq. (5) with appropriate substitutions.

Cross section σ_{H^+} for production of H^+ by H_2 primaries was determined by first calculating

$$\sigma_T=-\frac{1}{\alpha}\frac{1}{(N_{H_2}+N_H)}\frac{\Delta N_{H^+}}{\Delta p}, \qquad (11)$$

which represents the average cross section for H^+ production by both the H_2 molecules and H atoms in the primary beam. The cross section σ_{H^+} for H^+ production by the H_2 primaries alone and the cross section $\sigma_{0,1}$ for production of H^+ by H primaries alone are related to σ_T by the equation

$$\sigma_T=f_1\sigma_{0,1}+f_2\sigma_{H^+};\quad f_2=1-f_1, \qquad (12)$$

which can be rearranged in the form

$$\sigma_{H^+}=\sigma_T/f_2-\sigma_{0,1}f_1/f_2. \qquad (13)$$

All the quantities on the right of Eq. (13) are known from the measurements described. Values of $\sigma_{0,1}$ were taken from the present data shown on Fig. 2 using the cross section for H atoms of one-half the H_2 beam energy. Deduction of correct values of σ_{H^+} by use of this formula rests on the assumption that the H atoms produced by H_2^+ dissociation in the neutralizing chamber have the same electron loss cross section as H atoms produced by proton neutralization in the neutralizing chamber.

The results of the σ_{H^+} and $\sigma_{H_2^+}$ measurements are shown in Fig. 4. The uncertainties associated with counting statistics (standard deviations) are indicated by the error flags. These are believed to dominate all other sources of error. The relatively large uncertainties in the σ_{H^+} values are associated with the relatively

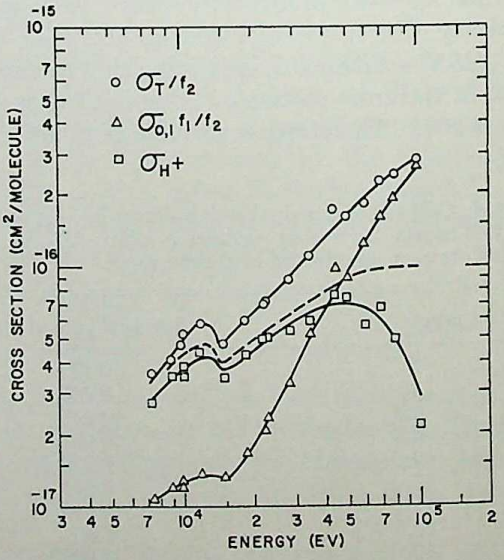


FIG. 5. Cross section σ_{H^+} for fast proton production by H_2 primaries and the two quantities σ_T/f_2 and $\sigma_{0,1}f_1/f_2$ [see Eq. (13)] which enter into the calculation of σ_{H^+}. The small difference between the upper curves at the higher energies leads to the large statistical uncertainties in σ_{H^+} at high energies. Dashed curve is the curve for σ_{H^+} obtained by using the lower values of $\sigma_{0,1}$ given by Curran and Donahue (Ref. 8) and Fogel (Ref. 7) instead of the $\sigma_{0,1}$ values from the present work in evaluating Eq. (13).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

small cross sections as well as the fact that σ_{H^+} is a difference between two quantities [Eq. (13)], each of which has associated statistical errors. This difference became quite small at high energies, leading to quite large percentage uncertainties. The two terms appearing on the right-hand side of Eq. (13) and the difference σ_{H^+} are plotted in Fig. 5 to show clearly the relative magnitude of the quantities as they occurred in the σ_{H^+} determinations.

The effect on σ_{H^+} of assuming the lower set of $\sigma_{0,1}$ values given by Curran and Donahue (4- to 40-keV H atom energy) and Fogel (40- to 50- keV H atom energy) is indicated by the dashed curve in Fig. 5. This curve is believed to represent a less reliable estimate of σ_{H^+} than the solid curve since the "low" values of $\sigma_{0,1}$ used in calculating the dashed curve may have resulted from calibration disparities which were eliminated in the present set of measurements by our use of a common basis of calibration in the $\sigma_{0,1}$ and σ_T measurements.

The cross section σ_H for fast H atom production in reactions 1 and 3 could not be determined accurately because the high flux of H atoms in the primary neutral beam tended to mask the production of secondary H atoms at small angles. However, an estimate of σ_H was made at 10 keV as follows. Careful measurements of the angular distribution of fast H atoms emergent from the collision chamber were made at two collision chamber pressures with results shown in Fig. 6. At angles greater than $\frac{1}{2}$ deg a very distinct pressure-dependent H atom component is present corresponding to an H atom production cross section of 1.25×10^{-16} cm². The portion of this cross section due to H_2 dissociation is estimated as 1.17×10^{-16} cm² by subtracting the calculated amount expected from Coulomb scattering of H atoms in the primary beam. By making a linear extrapolation of the large-angle H atom distributions to zero angle, the estimated ratio of total H atom production to production outside $\frac{1}{2}$ deg was found to be 2.14, giving a final estimate of 2.5×10^{-16} cm² for the value of σ_H.

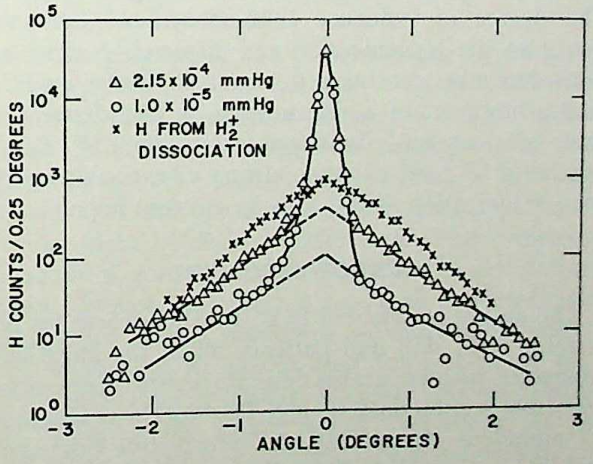


FIG. 6. Angular distribution of fast H atoms measured in order to estimate σ_H the cross section for production of fast H atoms from H_2 primaries. The large peak in the center is due principally to H atoms present in the primary beam and, consequently, shows no dependence on collision chamber pressure over the indicated pressure range 1.0×10^{-5} to 2.15×10^{-4} mm Hg. Dashed lines show extrapolations used in estimating the small-angle production of H atoms due to H_2 dissociation. Crosses indicate angular distribution of H atoms from H_2^+ dissociation in collision with H_2. All distributions are for 10 keV primaries and were obtained with a long slit (S_5) as the detector window. The distributions are, therefore, not the same as "differential angular distribution" in the usual sense of the term.

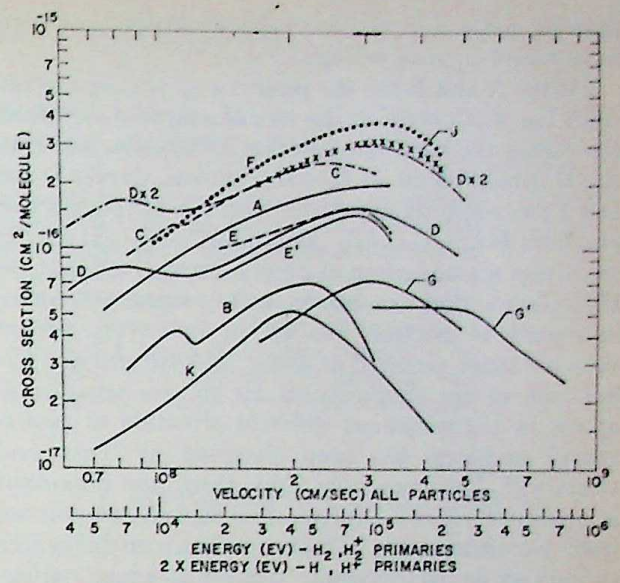


FIG. 7. Curve A: $\sigma_{H_2^+}$ the cross section for conversion of fast H_2 molecules to H_2^+ ions from Fig. 4. Curve B: σ_{H^+} the cross section for conversion of fast H_2 molecules to H^+ ions from Fig. 4. Curve C gives the sum of σ_{H^+} and $\sigma_{H_2^+}$ which equals the total cross section for electron loss from fast H_2 molecules. Curves D and D×2 are, respectively, 1 and 2 times the cross section $\sigma_{0,1}$ for electron loss from H atoms from Fig. 2. Curve E: ionization of H_2 by H primaries, including $\sigma_{1,-1}$ the ion production due to electron capture by the primary ions from Schwirtzke (Ref. 9). Curve E': curve E minus $\sigma_{1,-1}$ from Fig. 2. Curve F: cross section for free electron production in $H_2^+ + H_2$ collisions from Afrosimov (Ref. 15). Curve G and G' are, respectively, the results of Guidini (Ref. 16) and Sweetman (Ref. 2) for electron loss from H_2^+ ions in collision with H_2. Curve J: curve F with curve G subtracted to give cross section for electron removal from H_2 when struck by H_2^+. Curve K: cross section for production of slow H^+ ions in collision of H_2^+ with H_2 from Afrosimov (Ref. 15). A portion of curve D×2 between abscissa values 1.5 and 3×10^8 cm/sec was deliberately left out to avoid confusion with curve J.

The use of the linear extrapolation in estimating the small-angle H atom yield could not be rigorously justified, but seemed reasonable in view of the shape of the angular distribution of H atoms from H_2^+ primaries obtained with the same apparatus and shown superimposed in Fig. 6 for comparison.

DISCUSSION

No directly applicable theoretical calculations are available for comparison with the present experimental results; however, it is instructive to compare the previously available cross section data on several different hydrogen atom and molecule collisions with the new data on H_2+H_2 collisions. To facilitate such a com-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

parison, a number of cross section vs velocity curves have been compiled in Fig. 7.

Curves A and B are the present $\sigma_{H_2^+}$ and σ_{H^+} results from Fig. 4. Curve C is the sum of $\sigma_{H_2^+}$ and σ_{H^+} which represents the total cross section for electron loss from the H_2 projectile in an H_2+H_2 collision. Curves D and D×2 are, respectively, the cross section $\sigma_{0,1}$ for electron loss from H atoms and $\sigma_{0,1}$ multiplied by two. It may be seen from a comparison of curves C and D×2 that the H_2 molecule does not behave as two separate H atoms in regard to electron loss except in a very narrow velocity range centered at about 1.7×10^8 cm/sec. The question of the similarity of H_2 to two separate H atoms in the somewhat different situation of proton charge exchange has been discussed by Tuan and Gerjuoy.[14] They conclude that there are important interference effects between the capture amplitudes from two interacting H atoms which can make the capture cross section quite dependent upon nuclear separation. The above noted disparity between curves C and D×2 does not seem surprising in view of the possibility of analagous interference effects in the electron loss process. Curve E is the cross section given by Schwirtzke[9] for ionization of H_2 by H atoms while curve E′ is the same cross section with $\sigma_{1,-1}$ subtracted to eliminate the ionization contribution due to electron capture by the H primaries. It may be seen that curve C lies about a factor of 2 above curve E′, indicating that H_2 molecules are about twice as effective in removing electrons from H_2 molecules as are H atoms of the same velocity.

Curve F is the cross section for slow electron production in collisions of H_2^+ and H_2 as determined by Afrosimov *et al.*,[15] and curves G and G′ are, respectively, the results of Guidini[16] and Sweetman[2] for electron loss from fast H_2^+ ions in collisions with H_2. Curve J, the difference between curves F and G, represents the ionization of H_2 by H_2^+, not counting loss of electrons from the H_2^+ primaries. In the velocity range 1.7 to 2.2×10^8 cm/sec, curve J lies quite close to curve C, indicating that H_2^+ and H_2 are roughly equivalent in their power to remove electrons from H_2.[17] However, at 3×10^8 cm/sec there appears to be a real difference of about 50% between the two cross sections.

The cross section for removal of one electron from H_2 to form H_2^+ (curve A) is about three times larger than the cross section for removal of an electron from H_2^+ (curve G). This large difference seems reasonable since the first ionization potential of H_2 is 15 eV while the ionization potential of H_2^+ is 30 eV, indicating that the second electron is much more strongly bound than the first.

Curve K is the cross section for slow proton production in collisions of fast H_2^+ with H_2. This curve lies somewhat below curve B, indicating that the fast primary in H_2+H_2 collisions converts to H^+ with a slightly higher probability than the target particle in $H_2^++H_2$ collisions.

One of the more interesting results of this study is that the fast H atom yield from the pair H_2+H_2, measured at 10 keV, is only $\frac{1}{3}$ as great as the yield of fast H atoms from the pair $H_2^++H_2$ from Ref. 1. This may perhaps be explained by the fact that the lowest repulsive state of H_2 is the triplet state $1^3\Sigma_u$ which can only be reached from the ground state of H_2 by a spin flip or by electron exchange. Both of these processes are theoretically very unlikely compared with the simple near resonant process of electron transfer from an H_2 molecule to an H_2^+ ion, placing the newly formed molecule in a $1\,^3\Sigma_u$ repulsive state.

It is quite probable that the H_2 molecules which form the primary beam in our H_2 primary measurements (Fig. 4) are not in their ground vibrational states. This follows from the fact that the H_2^+ ion has a nuclear potential minimum at a nuclear spacing somewhat larger than that of the H_2 molecule. This has the effect of causing the H_2^+ ions to be formed preferentially in their third or fourth vibrational quantum level in which the vibrational excursion of nuclear spacing is quite large.[18] When the H_2^+ ions are subsequently converted back to H_2 by electron capture, the molecule may therefore be forced into very high vibrational levels. The degree of influence that vibrational excitation exerts on the H_2 ionization and dissociation cross sections has not yet been determined. Some evidence exists, however, of a dependence of H_2^+ dissociation cross sections on vibrational excitation.[1,19,20] Recent calculations[21] have shown a strong vibrational dependence of ionization of H_2^+ ions by electron impact.

ACKNOWLEDGMENT

The author is indebted to D. L. Allensworth for performing with skill and patience all of the hard labor connected with apparatus design, accelerator operation, and cross section data taking. The work was also aided by numerous stimulating and informative discussions with T. A. Green and J. M. Peek.

[14] T. F. Tuan and E. Gerjuoy, Phys. Rev. **117**, 756 (1960).
[15] V. V. Afrosimov, R. N. Ilin, and N. V. Fedorenko, Zh. Eksperim. i Teor. Fiz. **34**, 1398 (1958) [English transl.: Soviet Phys.—JETP **7**, 968 (1958)].
[16] J. Guidini, Compt. Rend. **253**, 829 (1961).
[17] Here we compare electron loss from the *target* H_2 molecule in the $H_2^++H_2$ collision with electron loss from the projectile in the H_2+H_2 collision.
[18] P. Marmet and T. Kerwin, Can. J. Phys. **38**, 972 (1960).
[19] A. C. Riviere and D. R. Sweetman, Proc. Phys. Soc. (London) **78**, 1215 (1961).
[20] C. F. Barnett and J. A. Ray, Third International Conference on Ionization Phenomena in Gases, 1963, University College, London (unpublished).
[21] J. M. Peek, Sixteenth Annual Gaseous Electronics Conference, 1963, Pittsburgh, Pennsylvania (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Bound States in a Debye-Hückel Potential

CHARLES RAY SMITH
National Bureau of Standards, Boulder, Colorado

(Received 13 June 1963; revised manuscript received 4 November 1963)

This is a study of some of the properties of the discrete energy levels of an electron in an exponentially shielded Coulomb potential, which is known in plasma physics as the Debye-Hückel potential and in nuclear physics as the Yukawa potential. A system with this potential possesses a finite number of bound states and these states are not degenerate with respect to the orbital angular momentum. First-order perturbation theory is used to obtain simple, analytical expressions for estimating the energy levels; the approximations appear to be quite accurate for large values of the Debye length. The perturbation calculations lead to an estimate of the number of bound states as a function of the Debye length. The matrix elements of the Hamiltonian for this potential are calculated in the representation of the hydrogen atom wave functions. The Hamiltonian matrix may serve as a convenient starting point for several other methods of computing the energy levels.

1. INTRODUCTION

ALTHOUGH the two-component gas with Coulomb interactions has not yet been adequately treated, certain results are available for various temperature and density regions. Specifically, the Debye-Hückel potential is reported to be the effective two-body interaction in the high-temperature, low-density region.[1] This result is often interpreted to mean that in a hydrogen plasma a bound electron moves in the potential given by the Debye-Hückel theory. By virtue of their spectra, atoms in a plasma serve as noninterfering probes, and a study of the effects of the Debye shielding on the energy levels of atomic hydrogen may be of more than academic value. For example, in the highly excited bound states of atoms in a plasma the Debye shielding may be approximately realized, and one wishes to examine the effects of this shielding on the atomic partition function.

It is the purpose of the present paper to estimate the energy levels of an electron in a Debye-Hückel potential (which, in nuclear physics, is known as the Yukawa potential). Although certain aspects of the problem have been investigated by others,[2–5] it is not apparent that the problem has been treated as generally as might be desired. Ecker and Weizel[2] restrict their study to s states and their method of calculation is open to question.[6] Harris[3] has used a variational technique for making numerical estimations of several energy levels, but it appears that her results are not useful for laboratory plasmas, since she chooses mostly small values for the Debye length. Her energy levels for the larger Debye lengths are in agreement with those calculated from analytical expressions obtained in the present paper.

[1] E. Meeron, J. Chem. Phys. 28, 630 (1958); E. W. Montroll, and J. C. Ward, Phys. Fluids 1, 55 (1958); F. Mohling, and W. T. Grandy, Jr., J. Math. Phys. (to be published).
[2] G. Ecker, and W. Weizel, Ann. Phys. 17, 126 (1956).
[3] G. M. Harris, Phys. Rev. 125, 1131 (1962).
[4] H. Margenau, and M. Lewis, Rev. Mod. Phys. 31, 594 (1959).
[5] H. W. Woolley, Air Force Special Weapons Center Tech. Doc. Rept. 62–21 (1962).
[6] O. Theimer, Z. Naturforsch 12a, 518 (1957).

2. PERTURBATION CALCULATIONS

The Debye (or Debye-Hückel) potential for an electron in the shielded field of a singly-charged ion is

$$V(r)=-\epsilon^2(e^{-r/d}/r). \tag{1}$$

Here ϵ is the unit of charge and d is the Debye length.

$$d=(kT/4\pi\epsilon^2 n)^{1/2}, \tag{2}$$

where k is the Boltzmann constant, T is the equilibrium temperature of the plasma, and n is the electron density. The potential in Eq. (1) possesses a finite number of energy levels, and these energy levels are not degenerate with respect to the angular momentum.

Atomic units are employed, so that the unit of distance is $\hbar^2/\mu\epsilon^2$ and the unit of energy is $-\mu\epsilon^4/2\hbar^2$, where μ is the reduced mass of the electron. Instead of the radial wave function $R(r)$, it is desirable to use $P(r)\equiv rR(r)$ so that the radial Schrödinger equation with the potential of Eq. (1) becomes

$$\frac{d^2P}{dr^2}+\left(\frac{2e^{-r/d}}{r}-W\right)P-\frac{l(l+1)}{r^2}P=0, \tag{3}$$

where l is the angular-momentum quantum number and W is the negative of the energy eigenvalue of the electron. Note that $W\geq 0$ for bound states. The Debye length for a hydrogen plasma is usually larger than 10^{-6} cm, which, in terms of atomic units, means $d\geq 10^3$.

The potential in Eq. (1) may be written as a Coulomb potential plus a perturbation $U(r)$:

$$V(r)=2/r+U(r)=2/r+2((e^{-r/d}-1)/r). \tag{4}$$

It is seen that the perturbing potential $U(r)$ becomes comparable to $V(r)$ when $r\sim d$. For the unperturbed case the average value, $\langle r\rangle$, of r goes as N^2, where N is the principal quantum number. Thus, it is expected that first-order perturbation theory will be accurate as long as $N^2\ll d$ and becomes progressively less accurate as N^2 approaches d. It is to be noted that the Debye potential is not meaningful for the lowest energy levels, since the electron orbits will not encompass

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

enough charged particles to realize any screening at all. Furthermore, for the lower states the orbital frequency of the electron may be large compared to the plasma frequency so that any static potential will be inapplicable to the problem.

The first-order energy shift ΔW_1 is given by

$$\Delta W_1=(P_{Nl},UP_{Nl})\equiv\int_0^\infty [P_{Nl}(r)]^2U(r)dr\,, \quad (5)$$

where the $P_{Nl}(r)$ are the normalized zero-order eigenfunctions of Eq. (3) and in this case are the radial hydrogen functions

$$P_{Nl}(r)=-\frac{1}{N}\left[\frac{(N-l-1)!}{[N+l)!]^3}\right]^{1/2}\times e^{-r/N}\left(\frac{2r}{N}\right)^{l+1}L_{N+l}^{2l+1}\left(\frac{2r}{N}\right), \quad (6)$$

where $L_{N+l}^{2l+1}(2r/N)$ is the associated Laguerre polynomial. The integration in Eq. (5) may be performed and yields

$$\Delta W_1=\frac{2^{2l+3}(N+l)!\,{}_2F_1(-N+l+1,\,-N+l+1,\,2l+2,\,4d^2/N^2)}{N^{2l+4}(2l+1)!(N-l-1)!(2/N+1/d)^{2N}d^{2N-2l-2}}-\frac{2}{N^2}, \quad (7)$$

where ${}_2F_1$ is the hypergeometric function.

If second and higher order terms in $1/d$ are neglected, Eq. (7) reduces to

$$\Delta W_1\cong -2/(d+N^2). \quad (8)$$

Hence, to first order in $1/d$, first-order perturbation theory provides the following expression for the energy levels of Eq. (3):

$$W\cong 1/N^2-2/(d+N^2). \quad (9)$$

Although Eq. (9) indicates that energy levels of Eq. (3) are not degenerate with respect to the angular momentum, it is to be noted that the degeneracy still exists if second and higher order terms in $1/d$ are neglected.

A useful approximation to ΔW_1 is obtained by expanding $U(r)$ in a power series in $1/d$. If this is done Eq. (5) becomes

$$\Delta W_1=-\frac{2}{d}+\frac{1}{d^2}\langle r\rangle-\frac{1}{3d^3}\langle r^2\rangle+\frac{1}{12d^4}\langle r^3\rangle-\frac{1}{60d^5}\langle r^4\rangle+\cdots. \quad (10)$$

For $P_{Nl}(r)$ in the form given in Eq. (6) Bethe and Salpeter[7] have an explicit formula for evaluating $\langle r^\nu\rangle$ through $\nu=4$. Equation (10) then becomes[8]

$$\Delta W_1\cong-\frac{2}{d}+\frac{1}{2d^2}[3N^2-l(l+1)]-\frac{N^2}{6d^3}[5N^2+1-3l(l+1)]+\frac{N^2}{96d^4}[35N^2(N^2-1)-30N^2(l+2)(l-1)+3(l+2)(l+1)l(l-1)]-\frac{N^4}{480d^5}[63N^4-35N^2(2l^2+2l-3)-5l(l+1)(3l^2+3l-10)+12]. \quad (11)$$

Once the $l=0$ values are obtained for a given N, a useful approximation for ΔW_1 with higher l values is

$$\Delta W_1=(P_{Nl},UP_{Nl})\cong(P_{N0},UP_{N0})-[l(l+1)/2d^3](d-N^2). \quad (12)$$

Again, it is noted from Eq. (12) that the angular momentum corrections are of order $1/d^2$ and higher.

Values of the perturbed eigenvalues W for a few values of N, l, and d are presented in Table I. The calculations are based on Eq. (11). (The lower N values are included only for completeness.)

If accurate estimates of the eigenvalues of the Debye potential become necessary, an efficient method for computing the eigenvalues is the Rayleigh-Ritz method.[9] This method involves diagonalizing the Hamiltonian matrix computed in some convenient representation. Good accuracy of this method depends on a judicious choice of base vectors, which, in this case, should be the unperturbed eigenfunctions given by Eq. (6). The matrix for the Hamiltonian for a given l (H is central

[7] H. A. Bethe, and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press Inc., New York, 1957), p. 17.

[8] H. W. Woolley (Ref. 5) obtained this expression through the d^{-4} term.

[9] J. C. Slater, *Quantum Theory of Atomic Structure* (McGraw-Hill Book Company, Inc., New York, 1960), Vol. I, Chap. 5.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Eigenvalues (W) as calculated from perturbation theory.

d (a.u.)	N	W $l=0$	$l=1$	$l=2$	Unperturbed
10^3	1	0.9980			1.0000
	2	0.2480	0.2480		0.2500
	3	0.1091	0.1091	0.1091	0.1111
	4	0.0605	0.0605	0.0605	0.0625
	5	0.0380	0.0380	0.0380	0.0400
	6	0.0258	0.0258	0.0258	0.0278
	7	0.0185	0.0185	0.0185	0.0204
	8	0.0137	0.0137	0.0137	0.0156
	9	0.0105	0.0105	0.0105	0.0123
	10	0.0081	0.0081	0.0081	0.0100
	11	0.0064	0.0064	0.0064	0.0083
	12	0.0051	0.0051	0.0051	0.0069
	13	0.0041	0.0041	0.0041	0.0059
10^2	1	0.9802			1.0000
	2	0.2306	0.2305		0.2500
	3	0.0924	0.0923	0.0921	0.1111
	4	0.0447	0.0446	0.0444	0.0625
	5	0.0233	0.0232	0.0230	0.0400

so that l values are not mixed) in this representation is

$$(P_{Nl},HP_{N'l})=(1/N^2)\delta_{NN'}+(P_{Nl},UP_{N'l}), \quad (13)$$

where

$$(P_{Nl},UP_{N'l})\equiv\int_0^\infty P_{Nl}2\left(\frac{e^{-r/d}-1}{r}\right)P_{N'l}dr. \quad (14)$$

With P_{Nl} given by Eq. (6) the matrix elements of $U(r)$, Eq. (14), become

$$A(N,N',l)\frac{\lambda_1^{\alpha+\alpha'-\gamma}}{(\lambda_1-k)^\alpha(\lambda_1-k')^{\alpha'}}\times{}_2F_1\left(\alpha,\alpha',\gamma,\frac{kk'}{(\lambda_1-k)(\lambda_1-k')}\right)$$
$$-A(N,N',l)\frac{\lambda_2^{\alpha+\alpha'-\gamma}}{(\lambda_2-k)^\alpha(\lambda_2-k')^{\alpha'}}\times{}_2F_1\left(\alpha,\alpha',\gamma,\frac{kk'}{(\lambda_2-k)(\lambda_2-k')}\right), \quad (15)$$

where

$$A(N,N',l)=\tfrac{1}{2}(2l+1)!\left(\frac{4}{NN'}\right)^{l+2}\times\left[\frac{(N+l)!(N'+l)!}{(N-l-1)!(N'-l-1)!}\right]^{1/2},$$

$$\gamma=2l+2, \qquad k=2/N, \quad k'=2/N',$$
$$\alpha=-N+l+1, \quad \lambda_1=1/N+1/N'+1/d,$$
$$\alpha'=-N'+l+1, \quad \lambda_2=1/N+1/N'.$$

The evaluation of these matrix elements is rather tedious but could be rapidly executed on an electronic computer.

3. THE NUMBER OF ENERGY LEVELS FOR A DEBYE-HÜCKEL POTENTIAL

If Eq. (9) were exact, then the principal quantum number would satisfy the condition

$$N^2\leq d \quad (16)$$

since it is required that $W\geq 0$ for bound states. In spite of the limited accuracy of first-order perturbation theory and the approximation made in arriving at Eq. (8), the result in Eq. (16) compares favorably with results of others.[2,4] In fact, Eq. (16) is identical to an expression obtained by Ecker and Weizel.

In connection with a hydrogen plasma, such an expression as Eq. (16) is of limited usefulness, since other perturbations in the plasma must also be considered,[10] especially for $N^2\sim d$. For example, if the l degeneracy is neglected, it is not to be assumed that $N=d^{1/2}$ energy levels must be included in the atomic partition function.

4. CONCLUDING REMARKS

The principal value of the results of this study is that several approximate expressions are made available, which reveal some of the interesting properties of the exponentially shielded Coulomb potential. It does not appear that these results alone have immediate application to the calculation of thermodynamic functions, since other large perturbative effects must also be superposed on the perfect gas behavior of a hydrogen plasma. At this point it is desirable to re-emphasize that the preceding calculations are relevant to plasmas only for large electron-protron separations and not for low-lying states. As already indicated, Eq. (16) should not be used in terminating the atomic partition function. In this connection it should be mentioned that the well-known work of Inglis and Teller,[10] which treats the effects of the interatomic Stark broadening, likewise should not be used in terminating the partition function. Their work is applicable only to estimating the number of distinct spectral lines emanating from a radiating gas: Even though the higher energy levels are Stark broadened to such an extent that the spectral lines coalesce into a continuum, these broadened levels still represent bound states and must be included in the atomic partition function. It is not at all clear how one should sum over such broadened levels.

[10] D. R. Inglis and E. Teller, Astrophys. J. 90, 439 (1939).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Slow Neutron Inelastic Scattering from Liquid Sodium*

P. D. RANDOLPH
Atomic Energy Division, Phillips Petroleum Company, Idaho Falls, Idaho
(Received 24 September 1963; revised manuscript received 7 February 1964)

Partial differential cross sections for slow neutron scattering from liquid sodium have been measured at 100, 150, and 200°C. The Materials Testing Reactor phased-chopper velocity selector was used with a 95% transmission sample at incident neutron energies of 0.025, 0.07, and 0.10 eV at eight scattering angles (16 to 85°). In the scattering law obtained from these cross sections, prominent coherent scattering maxima appear for energy transfers $\Delta E \leqq 0.2\ k_BT$, but disappear beyond $\Delta E = 0.3\ k_BT$. At each temperature, time-dependent widths $\rho(t)$ of the intermediate scattering function have been obtained by Fourier transformation of the scattering law for large momentum transfers. At times $<1\times10^{-13}$ sec these functions confirm the prediction of gas-like behavior, but at larger times are solid-like. The $\rho(t)$'s are also momentum-dependent at larger times indicating that the self-correlation function is not strictly Gaussian. Measurements of the zeroth and first moments of the energy transfer have been obtained from the scattering law at each temperature for several values of momentum transfer. The zeroth moment is in agreement with neutron diffraction experiments, but the first moment is as much as a factor of 2.5 larger than predicted theoretically.

I. INTRODUCTION

THIS report gives the results of an experimental measurement of the partial differential cross sections for slow neutron scattering by liquid sodium. Besides the cross sections themselves, which are of interest to reactor technology, information about the dynamics of the motions of atoms in the liquid can be obtained from them. Extensive investigations of this type have been made for water[1,2] and for liquid lead.[3] Sodium was chosen as the sample for this experiment because it meets the following requirements: (A) The liquid is monatomic so that no molecular effects occur. (B) The isotopic abundance is 100% Na^{23} so that the scattering from only a single isotope is involved. (C) It is one of the few elements meeting the first two requirements whose incoherent scattering cross section (1.85 b) is a large fraction of the total scattering cross section (3.4 b), and which also has a reasonably low absorption cross section (0.5 b). A relatively large incoherent cross section is desirable since it is only the incoherent scattering which yields information about the motions of individual atoms. A cold-neutron study of sodium has been made by Cocking[4] who studied the inelastic scattering in the range 80 to 200°C.

The present experiment was performed using the Materials Testing Reactor (MTR) slow-neutron phased-chopper velocity selector at incident neutron energies of 0.025, 0.07, and 0.10 eV at sample temperatures of 100, 150, and 200°C. The experimental data show that marked inelastic interference scattering occurs up to energy transfers $\Delta E \approx 0.2 k_BT$. Time-dependent width functions obtained from the data show that the self-part of the space-time correlation function is fitted well by the Gaussian approximation at very short time, but at intermediate times, Gaussian width functions differ by as much as a factor of 2 for different momentum transfers. We have also measured the first energy transfer moment of the scattering law at several fixed values of momentum transfer including corrections for multiple scattering and find that these are as much as a factor of 2.5 larger than predicted.

II. EXPERIMENTAL PROCEDURE

The MTR 4-rotor phased-chopper velocity selector used in this experiment has been described elsewhere in detail.[5] This velocity selector produces bursts of nearly monochromatic neutrons of thermal energies by the use of two Fermi choppers spinning at the same speed (5000 rpm) and phased with each other. The scattering target is placed 0.5 m beyond the second chopper and the scattered neutrons are detected simultaneously at eight scattering angles by BF_3 counter banks located 2 m from the sample. These detectors were at scattering angles of 16.3, 20.1, 26.0, 36.4, 47.6, 59.5, 72.1, and 84.7°, and subtended 7.5° in the scattering plane except the detector at 20.1° which subtended 3.5°.

The energy, flux, and time resolution of the incident neutrons are determined by a U^{235} fission chamber located in the transmitted beam 2 m from the sample. For each data run at a given temperature, the sample and empty sample are cycled alternately into the beam at 10-min intervals for a total running time of 48 h or longer. The data from each of the two holders are stored in a separate matrix of the 8192 channel time-of-flight analyzer. For this experiment, 256 10-μsec channels were used to time-analyze the data for each

* Work performed under the auspices of the U. S. Atomic Energy Commission.

[1] M. Sakamoto, B. N. Brockhouse, R. G. Johnson, N. K. Pope, J. Phys. Soc. Japan Suppl. BII, 370 (1962).

[2] K. E. Larson, N. Dahlberg, *Proceedings of the Chalk River Symposium on Inelastic Scattering of Neutrons in Solids and Liquids* (International Atomic Energy Association, Vienna (1963), Vol. 1, p. 317.

[3] B. N. Brockhouse, N. K. Pope, Phys. Rev. Letters 3, 259 (1959).

[4] S. J. Cocking, *Proceedings of the Chalk River Symposium on Inelastic Scattering of Neutrons in Solids and Liquids* (International Atomic Energy Agency, Vienna, 1963), Vol. 1, p. 227.

[5] R. M. Brugger, J. E. Evans, Nucl. Instr. Methods 12, 75 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

angle and also for the beam monitor. Background corrections are made by subtracting the empty-holder data from the filled sample-holder data. If necessary, these data are normalized to the same counting rate at long and short flight times. After corrections for detector efficiency, the data are then converted to cross section and scattering law. The data of this experiment comprise runs made at 100, 150, and 200°C at each of the incident energies of 0.025, 0.07, and 0.10 eV. The runs at 100 and 150°C were taken at the 8

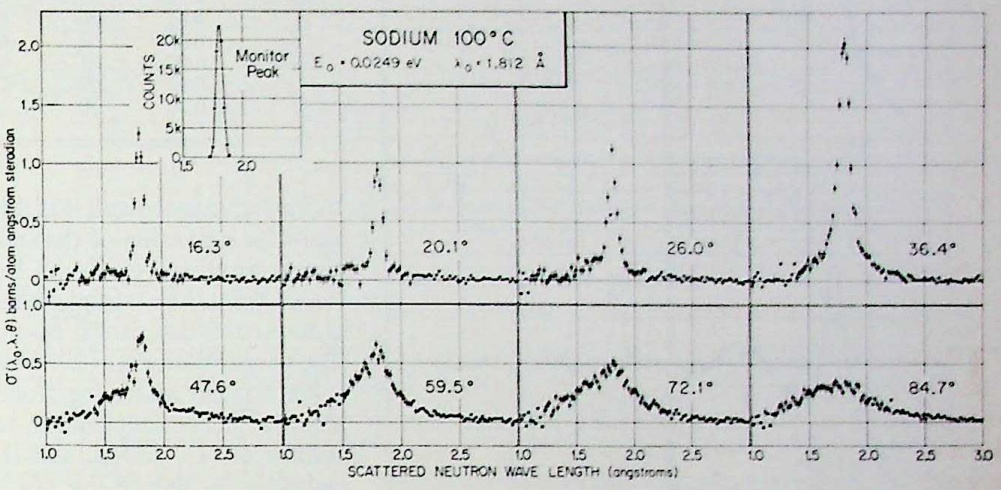


FIG. 1. Partial differential cross sections for scattering 0.0249 eV neutrons from liquid sodium at 100°C.

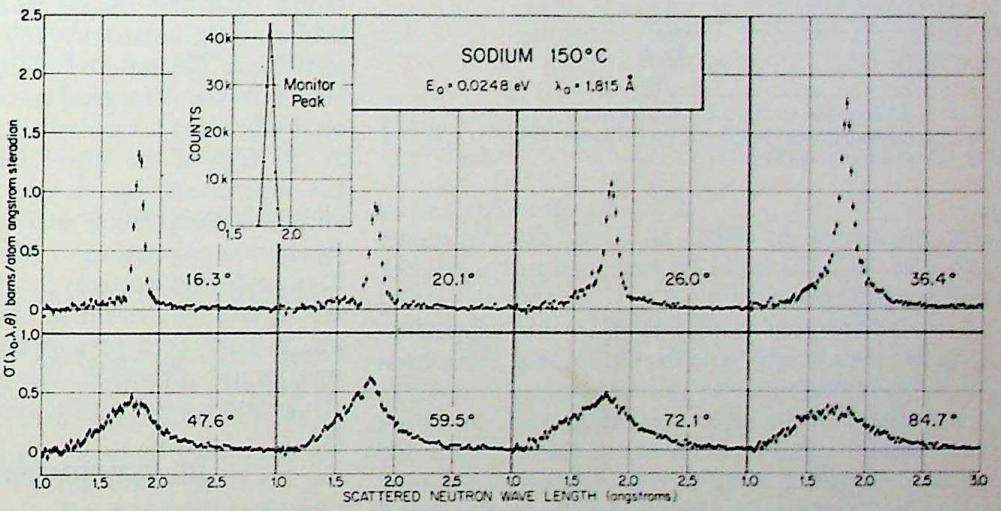


FIG. 2. Partial differential cross sections for scattering 0.0248 eV neutrons from liquid sodium at 150°C.

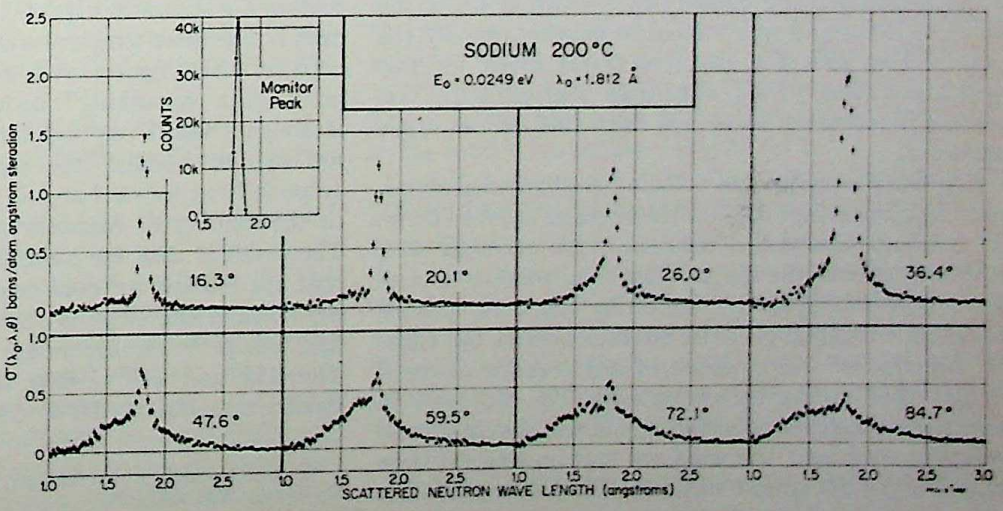


FIG. 3. Partial differential cross sections for scattering 0.0249 eV neutrons from liquid sodium at 200°C.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

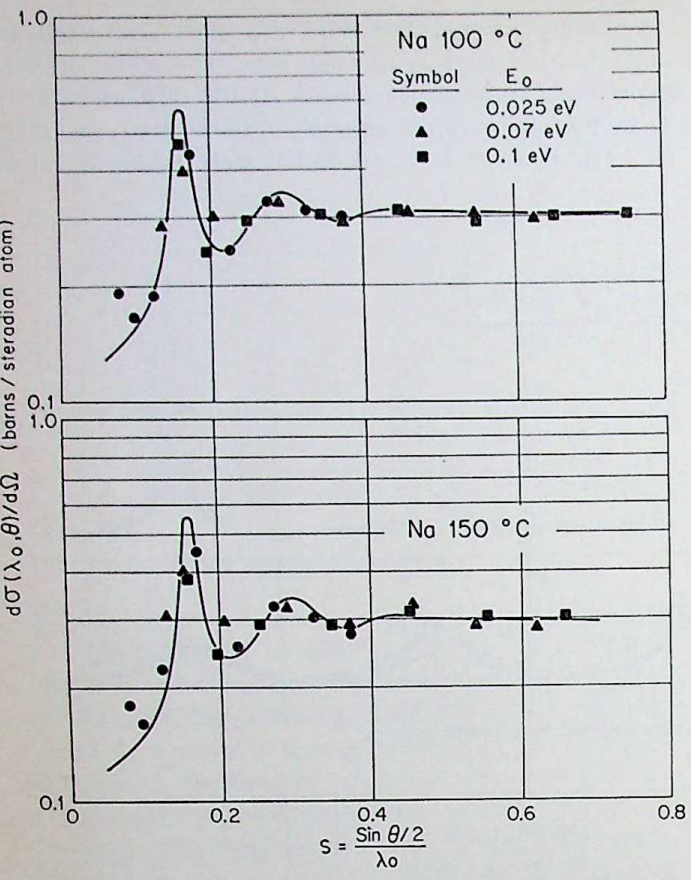


FIG. 4. Diffraction pattern for liquid sodium. The solid curve at both temperatures is the crystal spectrometer result at 100°C of Gingrich and Heaton (Ref. 7).

angles mentioned. At 200°C, additional detector angles of 84.4, 96, 107, 120, and 133° were used with a chopper speed of 6000 rpm.

The liquid-sodium sample used in this experiment was in the form of a rectangular slab having a thickness of 0.25 in. It was made as thin as possible in order to reduce effects of multiple scattering, and the above thickness is a compromise between this criterion and reasonable counting rates. The calculated transmission for thermal neutrons incident normally on the slab is 95%. In practice, the sample was oriented at 45° to the incident beam and was viewed in transmission by the detectors so that the effective transmission for this orientation was 91.5%. Multiple scattering in this sample is estimated to be less than 10% of the single scattering.

The sample holder was made of two hardened aluminum frames having a 0.25-in.-thick spacer between them. The frames sandwiched 0.001-in.-thick stainless steel windows against the spacer. The total window area of the holder was $5\frac{1}{2}\times 3$ in. Gasketing was accomplished by a narrow metal-to-metal contact around the edges of the window. No evidence of any leakage occurred at any time during the experiment. The difference in the thermal expansion coefficients of the aluminum and stainless steel kept the windows taut and no evidence of bulging of the sample was seen. The holder was made

TABLE I. Velocity selector energy resolution for elastic scattering.

Temp.	E_0(eV)	Δt(μsec)	$\Delta t/t$	$\Delta E/E_0$	$\Delta\beta$
100°C	0.025	36	0.0393	0.079	0.031
	0.07	36	0.0658	0.132	0.144
	0.10	36	0.0786	0.157	0.245
150°C	0.025	37.5	0.0410	0.082	0.028
	0.07	35	0.0143	0.129	0.124
	0.10	33.5	0.0735	0.147	0.201
200°C	0.025	31	0.0323	0.0645	0.020
	0.07	29.5	0.0567	0.113	0.097
	0.10	30	0.0657	0.131	0.160

taller than needed and was filled to less than capacity to allow for expansion of the sodium on melting. This further reduced any chance of bulging. The holder was not sealed from the atmosphere and in order to prevent excessive oxidation, it was operated under an argon purge. The purge volume was above the top of the beam so that there was no scattering from the argon. The window area was made large enough to accommodate the entire beam area (4×1.30 in.) and the frame was covered with cadmium to reduce the possibility of secondary scattering from the frame. Both the filled and empty holders were heated by 150-W cartridge heaters inserted into the top and bottom edges of each sample holder frame. The temperature was controlled automatically using a chromel-alumel thermocouple located at the edge of the spacer frame. After completion of the experiment, the sample holder was disassembled and the sodium sample analyzed. Spectrographic analysis showed only traces of Al, Ca, Cu, Fe, Mg, and Si. X-ray reflection patterns showed a thin film ($\sim$100 μ) of NaOH and NaO on the surfaces. The NaOH is believed not to have been present during the experiment but was formed afterward due to H_2O contamination of the kerosene used to store the disassembled sample.

III. EXPERIMENTAL RESULTS

Typical partial differential cross sections for liquid sodium are shown in Figs. 1–3 for an E_0 of 0.025 eV at each of the three temperatures. The counting rate data used for these figures were not smoothed. The "five-point data smoothing" technique[6] previously used, although it works very well for gases, is not satisfactory for sodium because the narrow quasielastic peaks (especially at forward angles) are comparable in width to the number of channels used to smooth the data. The result is that the smoothing gives counting rates that are too low (or even negative) at the base of the peak where the slope is changing rapidly. The error flags shown in the figures are counting statistics only. The 150 and 200°C data are of somewhat better quality than that at 100°C since the run times for these

[6] P. D. Randolph, R. M. Brugger, K. A. Strong, R. E. Schmunk, Phys. Rev. 124, 460 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

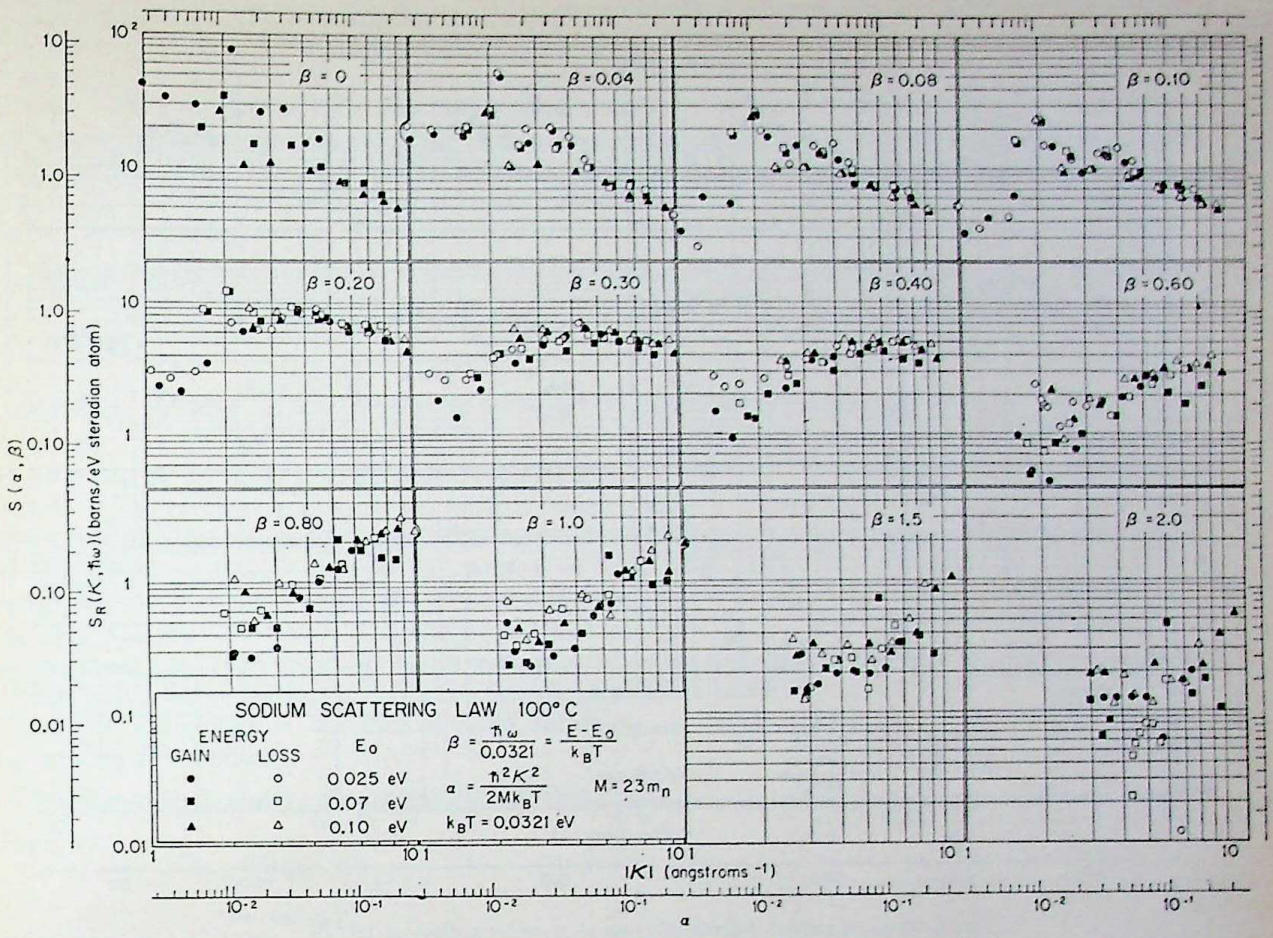


FIG. 5. Reduced partial differential cross sections for sodium at 100°C.

were each of 64 h, whereas the 100°C data were run only for 48 h.

A measure of the energy resolution for elastic scattering is given by the burst shapes of the incident beam which are also shown in the figures. Table I gives a summary of the widths of the bursts Δt (full width at half-maximum) measured by the beam monitor 2 m from the sample. This is calculated by $\Delta t/t_0$ where t_0 is the sample-to-detector time-of-flight at the incident energy E_0 and $\Delta E/E_0=2(\Delta t/t_0)$. The resolution is also expressed in terms of $\Delta\beta=\Delta E/2k_BT$, where T is the sample temperature.

The improvement in resolution at 200°C is due to the higher chopper speed (6000 rpm) used at this temperature.

The spectra for each temperature show two general features, a narrow quasielastic peak and a low broad spectrum of inelastically scattered neutrons. At the three forward angles, where the momentum transfers are small, the quasielastic scattering dominates and the inelastic scattering is very small. At 36°, the quasielastic peak is quite large compared with the other angles and the inelastic scattering has become larger. The momentum transfer for this quasielastic peak corresponds to the first maximum in the diffraction pattern so that this peak has a large contribution from coherent scattering. As the scattering angle continues to increase, the quasielastic peak becomes small or disappears and the inelastic scattering dominates the spectrum. The fact that at 200°C a quasielastic peak is seen at all angles is probably due to the improved resolution used at this temperature. Except for this, the spectra at each angle are qualitatively the same, and only small quantitative differences in the inelastic spectrum are present. At the 0.07 and 0.1 eV incident energies which are not shown, the data cannot be easily resolved into a quasielastic and an inelastic part. The cross sections shown in these figures are normalized to the total cross section value of 3.4 b.

Figure 4 shows the diffraction pattern obtained at 100 and 150°C by integrating the observed cross sections over the scattered wavelength. These are plotted as $d\sigma(\lambda_0)/d\Omega$ versus $s=\sin\frac{1}{2}\theta/\lambda_0$, and each plot includes the data from each of the three incident energies. The solid curve is the 100°C result of Gingrich and Heaton[7] obtained using a neutron crystal spectrometer and is normalized to the data at large s. The agreement is good except at the smaller values of s, where the finite

[7] N. S. Gingrich, L. Heaton, J. Chem. Phys. 34, 873 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

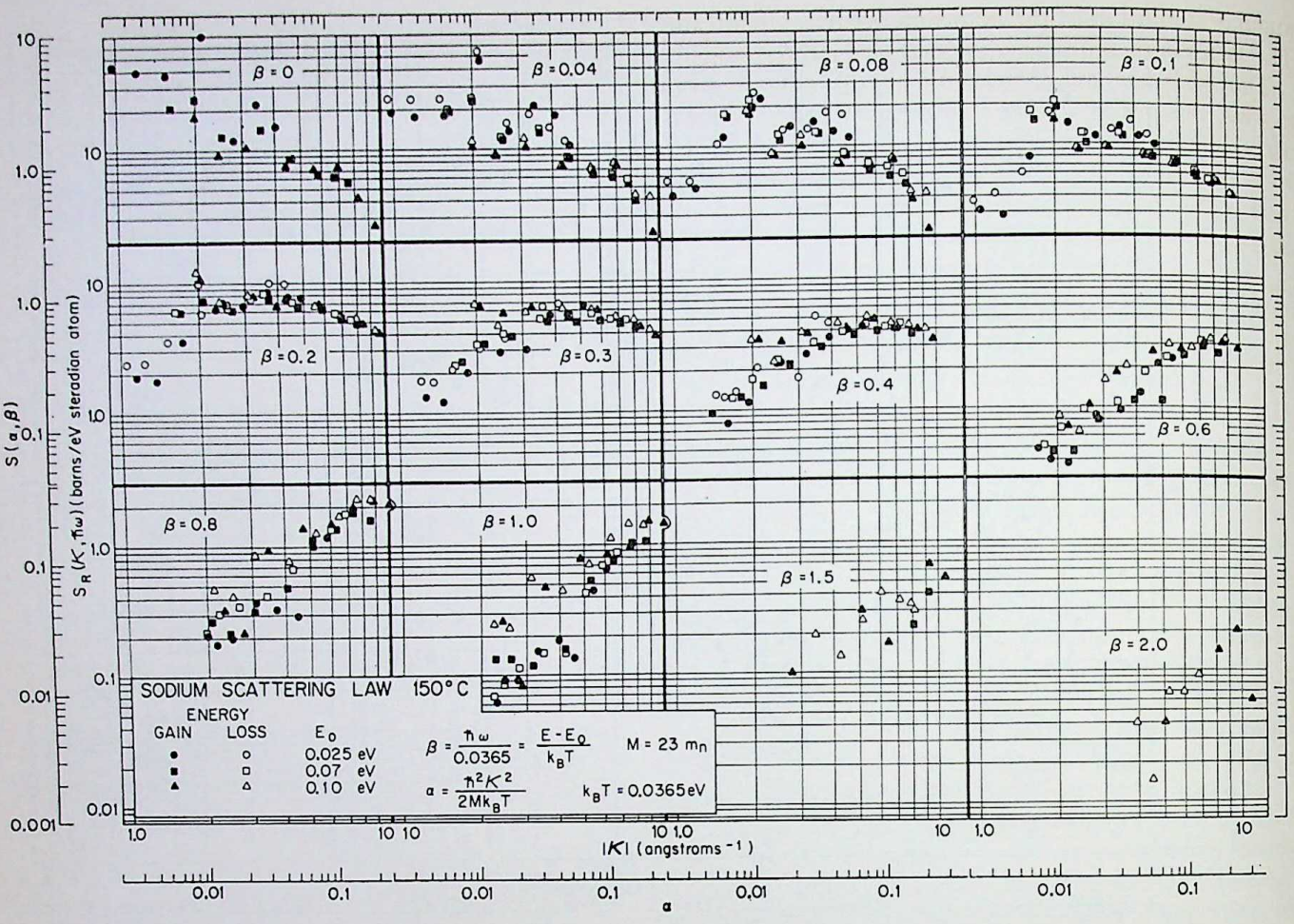


FIG. 6. Reduced partial differential cross sections for sodium at 150°C.

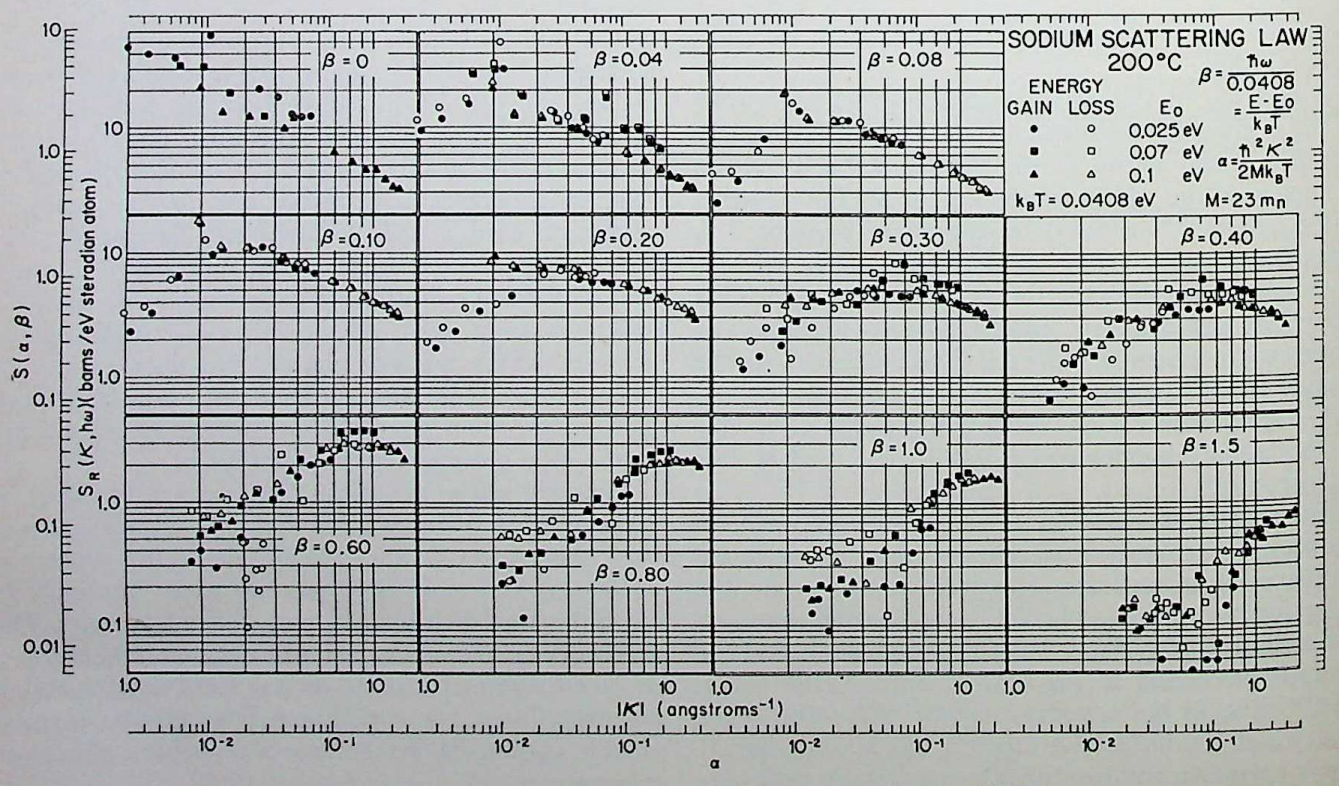


FIG. 7. Reduced partial differential cross sections for sodium at 200°C.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

angular resolution of the detectors becomes important. The two major diffraction maxima are clearly evident.

The complete presentation of all the data is given in Figs. 5–7 in terms of the scattering law. This form is used since it is representative of all the data and is a very compact form. The form of the scattering law used is the "reduced partial differential cross section[8]" $S_R(|\kappa|,\hbar\omega)$ and is defined by

$$d^2\sigma(E_0,E,\theta)/dEd\Omega = [E/E_0]^{1/2}\exp[-\hbar\omega/2k_BT]S_R(\kappa,\hbar\omega), \quad (1)$$

where

$$\hbar\omega = E - E_0,$$
$$\boldsymbol{\kappa} = \mathbf{k} - \mathbf{k}_0,$$

where E_0, E, $\mathbf{k}_0$, and $\mathbf{k}$ are the energy and reduced wave vectors of the incident and scattered neutron, respectively, and θ is the scattering angle. This form obeys the conditions of detailed balance and is even in the energy transfer and in the magnitude of the momentum transfer. The results at each temperature are plotted as S_R versus κ for fixed values of $\hbar\omega$ (in units of k_BT). This form has the dimensions of a cross section (b/eV sr atom) and is related to the dimensionless form[9] of the scattering law $S(\alpha,\beta)$ by

$$S_R(\kappa,\hbar\omega) = (\sigma_b/4\pi k_BT)S(\alpha,\beta), \quad (2)$$

where

$$\alpha = \hbar^2\kappa^2/2Mk_BT,$$
$$\beta = \hbar\omega/k_BT.$$

The mass of the scattering atom is given by M, T is the absolute temperature of the sample, and σ_b is the bound-atom cross section. Scales for $S(\alpha,\beta)$ are also shown in the figures.

The range of momentum transfers spanned at each incident energy overlap so that in these regions, the scattering law is overdetermined. At any fixed incident energy, the smallest scattering angle has the smallest κ value. At low-energy transfer ($\beta \lesssim 0.1$) the effects of finite energy resolution cause the data from the different incident energies to form separate curves at the smaller angles, the higher energy curves being below the lower. At larger angles where resolution effects are small, the different incident energies give a single curve. In spite of the energy resolution, the effects of interference scattering are clearly present. The interference maxima are seen, corresponding to those seen in Fig. 4. These maxima are present at each temperature out to energy transfer $\hbar\omega \lesssim 0.25\ k_BT$, and disappear beyond this. This fact demonstrates that in sodium, the convolution approximation[10] is not adequate for energy transfers larger than about $0.25\ k_BT$ since it predicts contributions to the coherent scattering that have magnitudes which are independent of the energy transfer. For small energy transfers at low-momentum transfers, however, cold-neutron results[4] show that the convolution approximation is in agreement with the experiment.

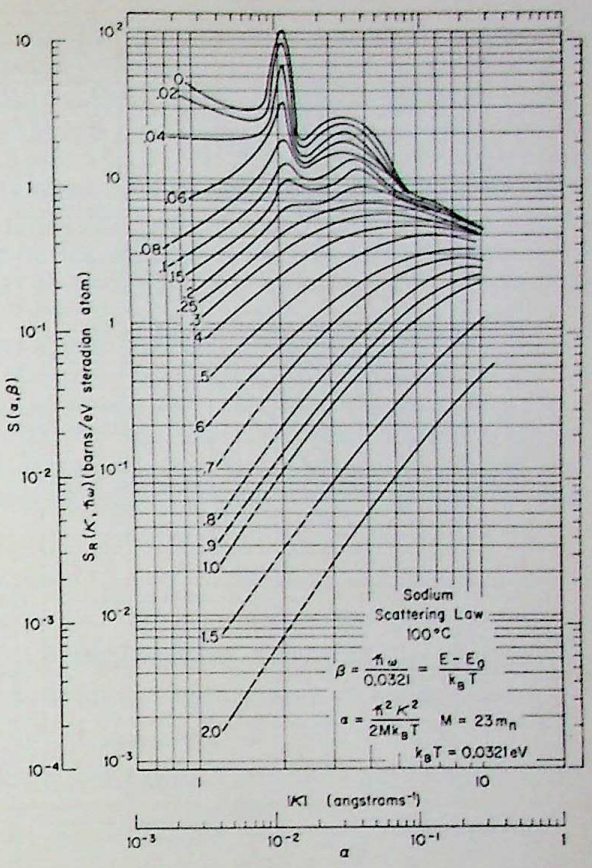


FIG. 8. Smooth curve scattering law for sodium at 100°C.

Smooth-curve scattering laws were obtained for each temperature by making a hand-drawn fit to the data. These are shown in Figs. 8–10. They are plotted as reduced partial differential cross sections $S_R(\kappa,\hbar\omega)$. Also shown are scales for the Egelstaff scattering law $S(\alpha,\beta)$. In drawing these, an attempt was made to reduce the effects of energy resolution by drawing the curve at low κ values for $\beta < 0.3$ to follow or lie above the lowest incident energy data.[11] Qualitatively, these curves are quite similar at each of the three temperatures indicating that from 100 to 200°C little change occurs in the properties of the liquid. All the smooth $S_R(\kappa,\hbar\omega)$ curves and the data plots are normalized to the total cross section using the value $\sigma = 3.4$ b.

IV. GAUSSIAN WIDTH FUNCTIONS

As was first shown by van Hove,[12] the scattering law is related to a time-dependent pair-correlation function for the system by a double Fourier transformation.

[8] R. M. Brugger, Atomic Energy Commission Research and Development Report No. IDO-16694 (Rev) 1962 (unpublished).
[9] P. A. Egelstaff, *Proceedings of the Vienna Symposium on Inelastic Scattering of Neutrons in Solids and Liquids* (International Atomic Energy Agency, Vienna 1961), p. 25.
[10] G. H. Vineyard, Phys. Rev. **110**, 999 (1958).
[11] R. M. Brugger, Phys. Rev. **126**, 29 (1962).
[12] L. van Hove, Phys. Rev. **95**, 249 (1954).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

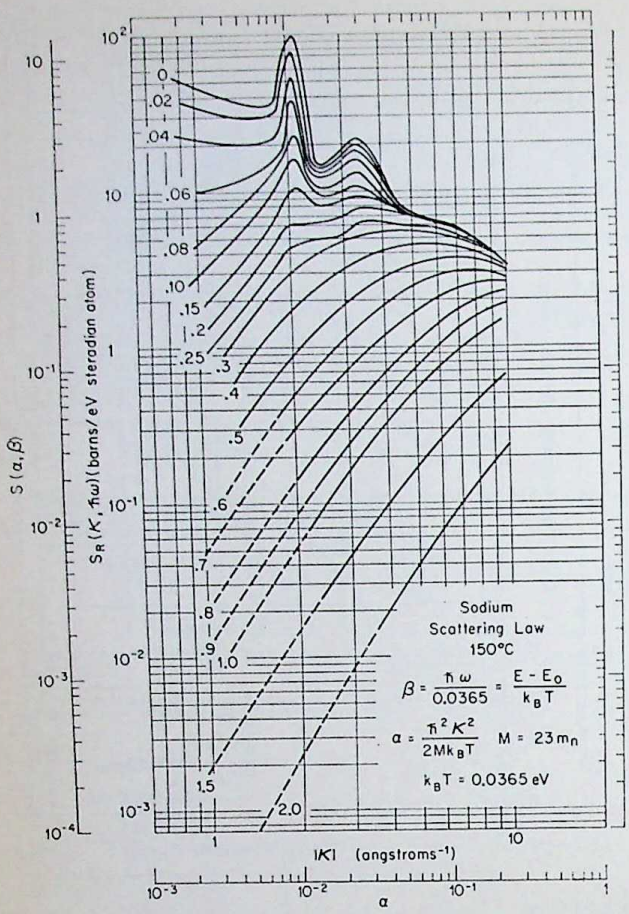


FIG. 9. Smooth curve scattering law for sodium at 150°C.

Various theoretical models for this correlation function have been studied for simple liquids. Vineyard[10] was the first to propose a specific model for the self-part of the correlation function which describes the self-diffusion of the liquid. The model proposed by Vineyard is that the self-part of the space-time correlation function is Gaussian in its space dependence with a time-dependent width. This model was shown to hold for a perfect gas, an oscillator, a Debye lattice, and for an atom obeying the classical equation of diffusion. This model is usually assumed to hold for simple liquids, at least as a first approximation. Indeed, it has been shown by Rahman *et. al.*,[13] that the Gaussian assumption enters as the first-order term in an expansion of the so-called intermediate scattering function in powers of κ^2.

In order to find out from experiment how well the Gaussian assumption holds, it is more convenient to work with this intermediate scattering function rather than with the correlation function or the scattering law. Defining a width function $\rho(t)$ by the Gaussian self-correlation function $G_s(r,t)$,

$$G_s(r,t)=[\pi\rho(t)]^{-3/2}\exp[-r^2/\rho(t)] \qquad (3)$$

then the intermediate scattering function $I_s(\kappa,t)$, which is the Fourier transform of G_s over space, may be written as a Fourier transform of the scattering law over the energy transfer.

$$I_s(\kappa,t)=\exp[-\kappa^2\rho(t)/4]=2\int_0^\infty \cos(\beta\tau)S_s(\alpha,\beta)d\beta, \qquad (4)$$

where in the right-hand member, we have used the symmetric form of the scattering law, and $\tau=(k_BT/\hbar)t$.

The use of the symmetric form of the scattering law implies that $\rho(0)=\hbar^2/2Mk_BT$ rather than $\rho(0)=0$. This constant term is of purely quantum-mechanical origin and is very small. From Eq. (4), if the Gaussian assumptions hold exactly, measurement of $S(\alpha,\beta)$ at fixed α (i.e., fixed κ^2) in principle allows a determination of $\rho(t)$ as was first pointed out by Vineyard.[10] Since $S(\alpha,\beta)$ is known for a range of α, then $\rho(t)$ is overdetermined. If the Gaussian assumption is exact, then the $\rho(t)$ obtained by Fourier transforming S at different values of α should give the same function and, hence, furnishes a test of the assumption.

Fourier transformation of the scattering law for liquid sodium has been performed at each of the three temperatures for several values of α using a program previously written.[11] Since it is only the noninterference portion of S that contains information about the diffusive motions, the range of α at which the trans-

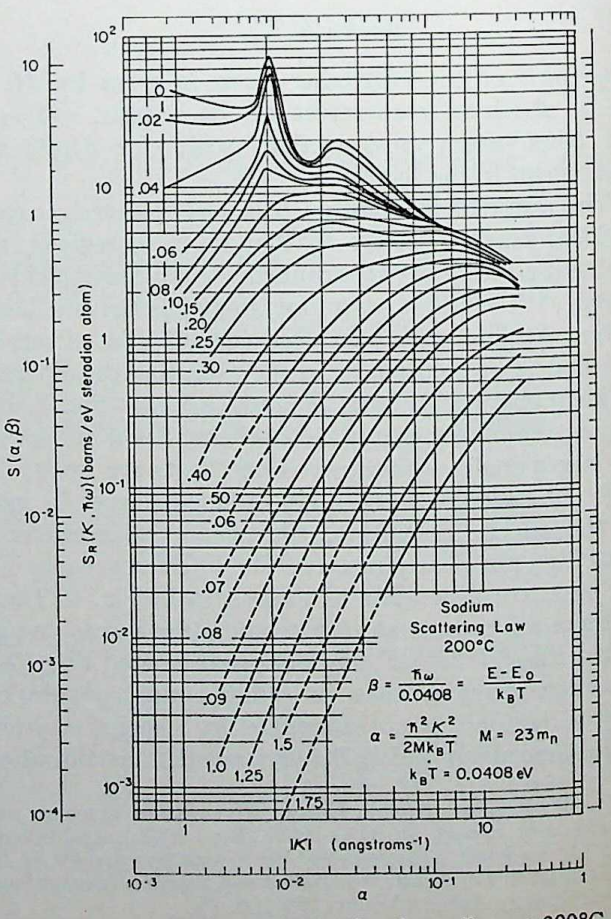


FIG. 10. Smooth curve scattering law for sodium at 200°C.

[13] A. Rahman, K. S. Singwi, A. Sjölander, Phys. Rev. 126, 986 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

formation was computed for $\alpha \geq 0.1$ at each temperature using the smooth curve scattering laws shown in Figs. 8–10. It is seen that for $\alpha \gtrsim 0.1$, the strong interference effects have disappeared. The advantage of this region of α space is that no approximate method need be used to separate the interference and the noninterference parts of the scattering. The disadvantage of this is that Fourier transformation at large α restricts the range for which $\rho(t)$ can be obtained to relatively short times. In making these transformations, $S(\alpha,\beta)$ was assumed to go linearly to zero beyond $\beta=2$. In all cases, this gave a zero value for $S(\alpha\beta,)$ at $\beta \lesssim 2.4$.

Figure 11 shows the resulting width functions plotted as a function of time for various fixed values of α. Also, there is shown for comparison, the slope characteristic of a simple diffusion model and the width functions for a perfect gas. Generally, the shapes of the various $\rho(t)$ are the same for each temperature. In none of the cases is $\rho(t)$ independent of α for $t \gtrsim 1\times10^{-13}$ sec, so that in this time range, the Gaussian assumption does not hold exactly. At a given time, the experimental $\rho(t)$ values drop by a factor of about 2 in the α range covered. For times less than 1×10^{-13} sec, the curves agree well and follow the gas behavior expected for very short times. The apparent large differences in the $\rho(t)$'s for $t \lesssim 0.1\times10^{-13}$ sec are due to the use of the logarithmic ordinate; on a linear plot these portions of the curves fall very close to each other and the average value of $\rho(0)$ is $\rho(0) \cong \hbar^2/2Mk_BT$. At larger times ($t > 1\times10^{-13}$ sec.), the slopes of the $\rho(t)$ curves are in all cases much less than that of the simple diffusion model. Assuming that the experimental $\rho(t)$ behavior is ap-

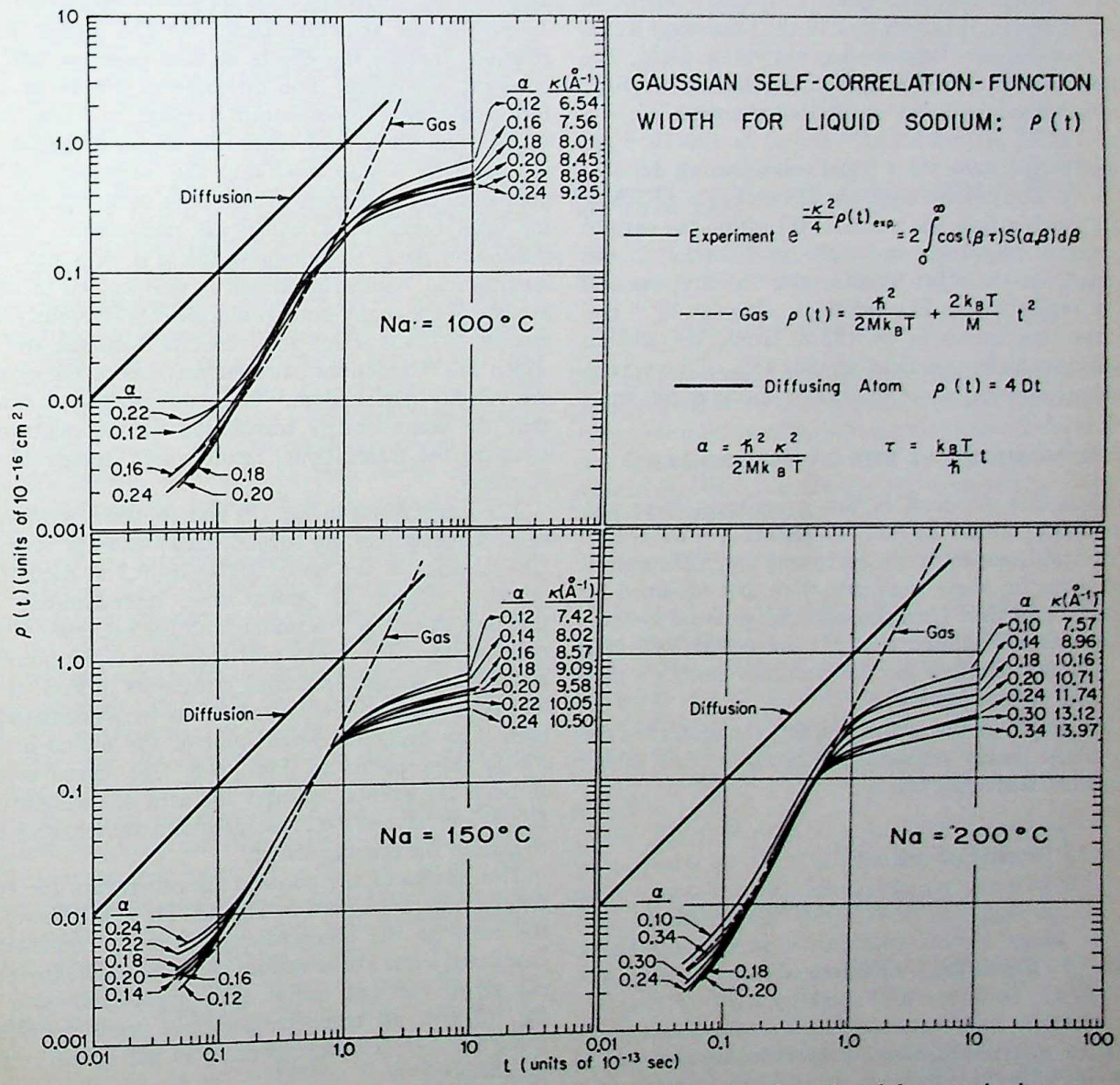


FIG. 11. Time-dependent width functions for sodium, obtained by Fourier transformation of the scattering law assuming a Gaussian form for the self-correlation function.

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

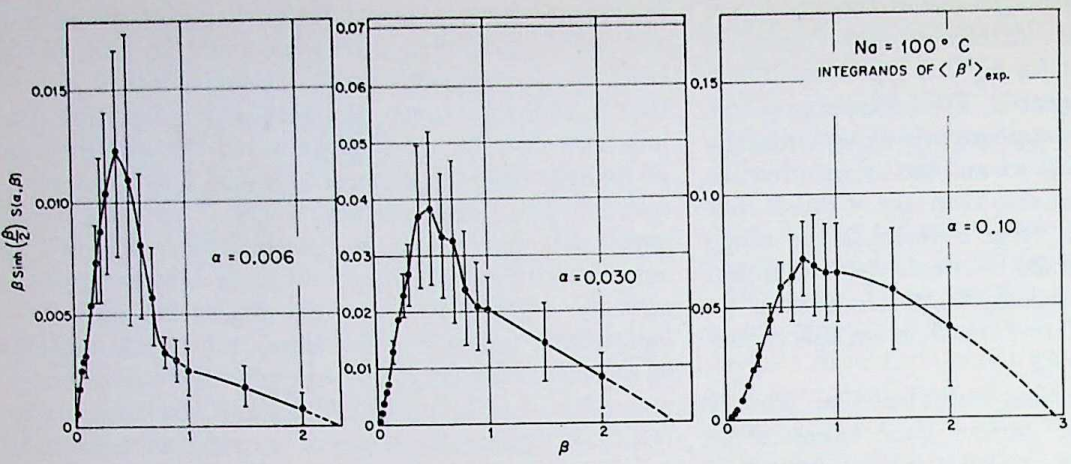


FIG. 12. Representative integrands of the first energy transfer moment of the scattering law for sodium at 100°C.

proximately Gaussian, and that at sufficiently long times it should be proportional to t, characteristic of simple diffusion, it is seen that in the time range available, such diffusive behavior has not yet occurred. The flatness of the curves indicates that for $t<10\times10^{-13}$ sec, the behavior is more nearly that characteristic of a solid. Taking $\rho(t)=0.5\times10^{-16}$ cm^2 at $t=10\times10^{-13}$ sec then the rms value of r obtained assuming Eq. (3) is 0.87 Å. The nearest neighbor distance r_0 is 3.83 Å so that at such a time the atom is still within the cage of its nearest neighbors and diffusive behavior is not expected. On the other hand at $t\cong0.8\times10^{-13}$, the end of the region of gas-like behavior, $\rho(t)=2\times10^{-16}$ cm^2 and the rms radius is $r=0.55$ Å. Thus, the sodium atoms move as free particles for only 15% of the nearest neighbor distance before feeling the effects of binding.

V. MOMENTS OF THE SCATTERING LAW

The results obtained in this experiment have also been used to obtain the first two moments of the energy transfer with respect to the scattering law.[14] Theoretical expressions for these moments were first obtained by Placzek[15] and have been discussed for classical systems by DeGennes.[16] Rahman *et. al.*,[13] have recently rederived the moment relations for the quantum case in a very clear and concise fashion.

Letting $\langle\beta^0\rangle$ be the zeroth and $\langle\beta^1\rangle$ the first, then the two lowest energy moments are given in terms of the symmetric scattering law by

$$\langle\beta^0\rangle\equiv2\int_0^\infty \cosh(\tfrac{1}{2}\beta)S(\alpha,\beta)d\beta = (1/\sigma_b)\{\sigma_{\rm inc}+\sigma_{\rm coh}[1+\gamma(\kappa)]\}\,, \quad (5)$$

$$\langle\beta^1\rangle\equiv2\int_0^\infty \beta\sinh(\tfrac{1}{2}\beta)S(\alpha,\beta)d\beta=\alpha\,, \quad (6)$$

where σ_b, $\sigma_{\rm inc}$ and $\sigma_{\rm coh}$ are the total bound, incoherent, and coherent scattering cross sections, respectively, and $1+\gamma(\kappa)$ is the structure factor for the liquid. These relations include the effects of both coherent and incoherent scattering. The integrations are to be performed at constant momentum transfer (i.e., constant α) rather than for a fixed scattering angle. Equation (5) is essentially a statement that the integration of the double differential cross section over outgoing energies for fixed momentum transfer is proportional to the diffraction pattern. Equation (6) is a sum rule, the longitudinal f-sum rule which is well known in other aspects of the many-body problem. It is rigorously correct for systems which are reflection invariant and for which the interactions between particles of the system are velocity-independent.[17] Roughly speaking, it states that the mean energy transferred by the neutron is equal to the translational recoil kinetic energy of the scatterer.

The integrations in Eq. (5) and (6) were performed, using a planimeter, for several fixed values of momentum transfer in the range $0.004\leq\alpha\leq0.2$ at each temperature. Figure 12 shows some representative integrands of $\langle\beta^1\rangle$ at the middle and the extremes of the range of logα. The values of $S(\alpha,\beta)$ were taken from the normalized smooth scattering law curves of Figs. 8–10. The error flags on the plotted points were estimated at each β by using the lower limit of the spread of the $S(\alpha,\beta)$ data points of Figs. 6–8. The curve drawn through the points indicates the area of intergration. In all cases beyond $\beta=2$, the integrand was extrapolated to zero in the manner shown.

The results of the moment measurements for each temperature are shown in Fig. 13. These are shown as the ratio of the observed experimental moments to calculated value at the various values of α. The theoretical zeroth-moment ratios [the right-hand side of Eq. (5)] at all three temperatures were calculated using for $1+\gamma(\kappa)$ the experimental diffraction results

[14] A brief report on the results of this section has been given in Phys. Letters 3, 162 (1963).
[15] G. Placzek, Phys. Rev. 86, 377 (1952).
[16] P. D. deGennes, Physica 25, 825 (1959).
[17] D. Pines, in *The Many Body Problem*, edited by D. Pines (W. A. Benjamin Inc., New York, 1961), p. 39.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Ratio of observed-to-theoretical moments of the energy transfer.

Temperature	100°C		150°C		200°C	
$[\langle\beta^0\rangle_{obs}/\langle\beta^0\rangle_{theor}]_{av}$	1.06		1.14		1.11	
$[\langle\beta'\rangle_{obs}/\alpha]_{av}$	2.76	2.30[a]	2.25	1.82[a]	1.66	1.42[a]
$[\langle\beta'\rangle_{obs}/\alpha]_{max}$	2.05	2.64[a]	2.75	2.22[a]	1.92	1.80[a]

[a] Values corrected for multiple scattering.

of Gingrich and Heaton[7] at 100°C. The results for the zeroth-moment ratios cluster near the value 1 in agreement with Eq. (5). The measured first moments, however, are in marked disagreement with the expected values, being as much as a factor 3 larger than the calculated value. In addition, the discrepancy is temperature-dependent being largest at the lowest temperature. At each temperature, the first-moment ratio also exhibits an apparent α dependence, the ratio decreasing at large and small α. The decrease at large α would be expected since for sufficiently large momentum transfers, the system would act like an ideal gas, and hence, would obey the sum rule. Also, the experimental trend indicates that for α sufficiently small, the sum rule may also be obeyed. In Table II are shown the mean values of the moment ratios plotted in Fig. 13. Also shown are the maximum values of the first moment ratio.

The errors on the first-moment ratio values in Fig. 13 are estimated to be $\pm 35\%$ as obtained from the spread in the scattering-law data points at the various beta values. It is seen that even on the low-limit side of the errors, the first-moment ratios would still disagree with the predicted values. The two major sources of systematic error which could lead to high values of $\langle\beta^1\rangle_{obs}$ are the finite energy resolution of the velocity selector and multiple inelastic scattering in the target. Qualitative corrections for finite resolution, which were made when drawing the scattering law, were checked at 100°C by convoluting the known velocity-selector burst shape with the partial differential cross section regenerated from this smooth-curve scattering law at the smallest scattering angle (16.3°) and the highest incident energy (0.1 eV), where resolution effects are the largest. The resulting cross section agrees within $\pm 10\%$ with data for that angle. Also, the effects of energy resolution are in general most pronounced for the smallest energy transfers ($\beta \leq 0.2$), and this region contributes little to the first moment integral (see Fig. 12). In any case, at large scattering angles (large alpha) where the cross section peaks are broad compared with the burst width, energy resolution effects would be small.

Since the target is thin, scattering only 8% of the incident beam, multiple scattering effects should be small. In order to determine to what extent multiple scattering affects the first moment, an estimate of the multiple scattering was made by fitting $S(\alpha,\beta)$ to a polynomial in α and assuming that the zeroth-order

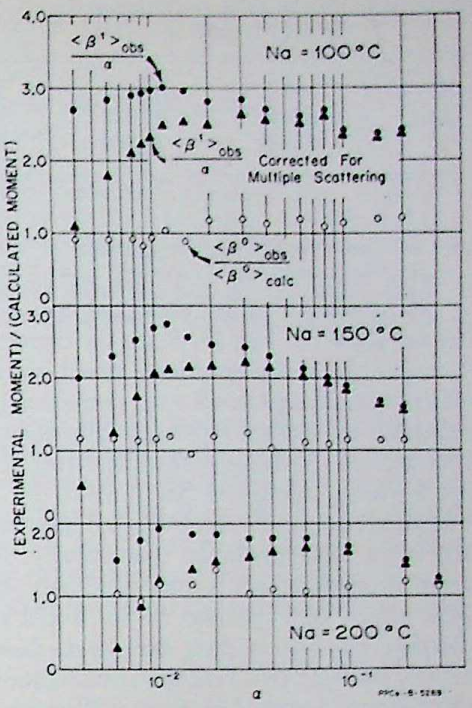


FIG. 13. Ratios of the observed to the calculated zeroth and first energy transfer moments of the scattering law for sodium at 100, 150 and 200°C.

term arises from multiple scattering,[1] i.e.,

$$S(\alpha,\beta)=A_0(\beta)+A_1(\beta)\alpha+A_2(\beta)\alpha^2+A_3(\beta)\alpha^3+\cdots.$$

Since the first term in a theoretical expression of $S(\alpha,\beta)$ is proportional to α and since multiple scattering is expected to be roughly isotropic and hence independent of α, it is assumed that the term $A_0(\beta)$ is due to multiple scattering. The first moment ratio, corrected for multiple scattering is then given by

$$\langle\beta^1\rangle_{corr}/\alpha=(\langle\beta^1\rangle_{obs}/\alpha)-(\langle\beta^1{}_c\rangle/\alpha),$$

where the correction term $\langle\beta_c{}^1\rangle$ is obtained from the same integral as Eq. (6) with $S(\alpha,\beta)$ replaced by $A_0(\beta)$. This correction is not expected to be exact, but the procedure is a reasonable one and is the best practical one available at the present time. Its accuracy is unknown and cannot be determined until experimental measurements of multiple inelastic scattering are available.

The $S(\alpha,\beta)$ curves of Figs. 8–10 for $\beta \geq 0.25$ were least-squares fit to a third-order polynomial in α in order to obtain $A_0(\beta)$. Scattering-law values for $\beta<0.25$ were not used since the extrema in $S(\alpha,\beta)$ caused by coherent scattering would necessitate a much higher order polynomial. The values of $A_0(\beta)$ for $\beta<0.25$ were obtained by extrapolating $A_0(\beta)$ to $\beta=0$. Table III shows the amount of multiple scattering compared to the peak value of $S(\alpha,\beta)$, i.e., $A_0(0)/S(\alpha,0)$ at each temperature together with the value of $\langle\beta_c{}^1\rangle$, the correction to the observed first moment. Table II gives

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Multiple scattering correction to the observed first energy transfer moment.

Temperature °C	100	150	200
$A_0(0)/S(\alpha,0)$	0.027	0.059	0.092
$\langle\beta_c^1\rangle$	0.0064	0.0063	0.0072

the values of the average and maximum of the first-moment ratio, corrected for multiple scattering. The resulting first-moment ratios corrected for multiple scattering are shown in Fig. 13. Since the correction to the first-moment ratio is proportional to α^{-1}, the largest effect occurs at small α. At small α the correction is large and accounts for most of the measured ratio and may be an over correction. At large momentum transfers the correction becomes negligible. In general, for alpha $\gtrsim 0.010$ at all temperatures the corrected moments are still much larger than predicted theoretically. A conceivable explanation for this is that velocity-dependent interactions might be important in the liquid or else that reflection invariance does not hold. However, it is extremely unlikely that reflection invariance would fail in a monatomic liquid. Also there is little reason to expect, nor is there other evidence to support the hypothesis that velocity-dependent interactions play any great role in a liquid. Additional scattering experiments of higher precision would be needed to make such hypotheses tenable.

SUMMARY AND CONCLUSIONS

The scattering law for liquid sodium between 100 and 200°C has been measured in the range of momentum transfer $1 \lesssim \kappa \lesssim 10$ Å^{-1} for energy transfers out to about 0.06 eV. The scattering law shows pronounced maxima due to inelastic coherent scattering for energy transfers below 0.2 k_BT, but which disappear for larger energy transfer, demonstrating that the convolution approximation does not hold in general for liquid Na. Time-dependent Gaussian width functions obtained by Fourier transformation of the scattering law at several momentum transfers show that sodium behaves as an ideal gas for times less than 1.0×10^{-13} sec corresponding to an rms atomic displacement of 15% of the nearest neighbor distance, but for larger times behaves more like a solid. In the time range available, $\lesssim 10\times10^{-13}$ sec, there is no evidence that simple diffusive behavior has been approached. The width functions for times greater than 1×10^{-13} sec are momentum-dependent, showing that the Gaussian assumption for the self-correlation function does not hold rigorously for momentum transfers larger than $K \approx 6$ Å^{-1}. Measurements of the two lowest moments of the energy transfer show that the zeroth moment gives the expected results but that the first moment is much larger than the theoretical values for a wide range of momentum transfers. The cause of this discrepancy remains to be explained. Whatever the underlying physical reasons for the discrepancy, it should not be peculiar to sodium, but should also appear in other metals. In any case, it would be of interest if energy transfer moments were measured in other metallic liquids.

ACKNOWLEDGMENTS

The author is much indebted to Dr. G. W. Griffing and G. Marshall for many helpful discussions, and to Dr. R. M. Brugger and Dr. R. G. Fluharty for advice and encouragement during this experiment.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Theory of Solid Neon at 0°K

WILLIAM J. MULLIN
Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio
(Received 12 November 1963)

A quantum-mechanical variational calculation of the energy per particle, the equilibrium nearest-neighbor distance, and the pressure-volume relationship for solid neon at 0°K is performed. Use is made of a Jastrow-type trial wave function, a phenomenological Mie-Lennard-Jones 12-6 potential, and a cluster expansion evaluation of the energy expectation value. It is found that correlation effects are negligible at equilibrium, contrary to the speculations of some authors, and that good agreement with the experimental data can be achieved if one of the interparticle potential parameters, ϵ, is changed by a small amount from its accepted gas-phase value. This parameter change is discussed.

I. INTRODUCTION

THE atoms of the inert gases have closed electronic shells which minimizes the importance of the details of their atomic structure in any discussion of the solid phase of these elements. The fact that only the gross properties are significant makes these elements ideal for testing basic principles of quantum mechanics and solid-state theory. Despite the straightforwardness of theoretical treatment, some discrepancies have been found in comparison of theory and experiment. For one, it is found that the predicted cohesive energy of solid neon at 0°K is too small.[1,2] (One writer[3] has predicted a cohesive energy which is in agreement with experiment; we will discuss this result in Sec. IV.)

Some authors[1,2] have speculated that the disagreement of the theoretical and measured values is caused by correlation effects. We will study this conjecture by using a Jastrow-type trial wave function[4] in a variational calculation. A cluster expansion will be used to evaluate the energy expectation value. The Mie–Lennard-Jones 12-6 potential is used to describe phenomenologically the electron interaction energy. We will show that correlation effects are negligible in solid neon near the equilibrium density and that theory and experiment can be brought into reasonable agreement by altering a Mie–Lennard-Jones potential parameter value by a small amount.

II. THE SCHRÖDINGER EQUATION AND THE TRIAL WAVE FUNCTION

The Schrödinger equation for both electrons and nuclei of the solid is

$$H\Psi(\mathbf{x},\mathbf{x}_e)=E\Psi(\mathbf{x},\mathbf{x}_e)\,, \tag{1}$$

where

$$H=H(\mathbf{x},\mathbf{p},\mathbf{x}_e,\mathbf{p}_e)=(1/2m)\sum_i \mathbf{p}_i^2+(1/2m_e)\sum_j \mathbf{p}_{ej}^2 +U(\mathbf{x},\mathbf{x}_e)\,, \tag{2}$$

with m, $\mathbf{x}_i$, $\mathbf{p}_i$, the mass, position vector, and momentum operator of the ith nucleus, respectively; m_e, $\mathbf{x}_{ej}$, $\mathbf{p}_{ej}$, the mass, position vector, and momentum operator of the jth electron. $U(\mathbf{x},\mathbf{x}_e)$ is the total electrostatic energy of the system, with the symbol $\mathbf{x}$, for example, representing the set of vectors $\mathbf{x}_1$, $\mathbf{x}_2$, $\cdots$. The Born-Oppenheimer approximation[5] can be applied to Eq. (1) to yield an equation for the nuclear motion alone,

$$[(1/2m)\sum_i \mathbf{p}_i^2+V(\mathbf{x})]\psi(\mathbf{x})=E\psi(\mathbf{x})\,. \tag{3}$$

The wave function $\psi(\mathbf{x})$ is that of the nuclei; $V(\mathbf{x})$ is the electronic energy.

The most widely used form for $V(\mathbf{x})$ is the phenomenological Mie–Lennard-Jones 12-6 potential

$$V(\mathbf{x})=4\epsilon\sum_{i<j}[(\sigma/x_{ij})^{12}-(\sigma/x_{ij})^{6}]\,, \tag{4}$$

where $x_{ij}=|\mathbf{x}_i-\mathbf{x}_j|$. Gas data are normally used to evaluate σ and ϵ.

If we now let

$$\lambda^2=\hbar^2/(2m\sigma^2\epsilon)\,, \tag{5}$$

then the combination of Eqs. (3) and (4) gives

$$[-\lambda^2\sum_i \nabla_i^2+4\sum_{i<j}(x_{ij}^{-12}-x_{ij}^{-6})]\psi(\mathbf{x})=E\psi(\mathbf{x})\,, \tag{6}$$

where x_{ij} and E are now dimensionless, i.e., we have made the substitutions $x_{ij}/\sigma\rightarrow x_{ij}$ and $E/\epsilon\rightarrow E$.

Our procedure is to solve Eq. (6) by means of a variational technique. Because of the large repulsive term, x_{ij}^{-12}, in the potential (4), a wave function for the solid must tend to zero when any two nuclei approach closely. One way to provide for this is to approximate the total wave function $\psi(\mathbf{x})$ by a product of single-particle, localized, nonoverlapping wave functions,

$$\psi(\mathbf{x})=\prod_{i=1}^{N}\phi(\mathbf{x}_i-\mathbf{R}_i)\,. \tag{7}$$

$\mathbf{R}_i$ is the position of the ith lattice site. Bernardes[1,3] chose $\phi(\mathbf{x}_i-\mathbf{R}_i)$ to be the wave function of a particle moving freely in a sphere centered at $\mathbf{R}_i$. The radius

[1] N. Bernardes, Phys. Rev. 112, 1534 (1958).
[2] L. H. Nosanow and G. L. Shaw, Phys. Rev. 128, 546 (1962).
[3] N. Bernardes, Phys. Rev. 120, 807 (1960).
[4] R. Jastrow, Phys. Rev. 98, 1479 (1955).
[5] M. Born and R. Oppenheimer, Ann. Phys. 84, 457 (1927).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the sphere was used as a variational parameter. In this theory an atom feels a force exerted on it by each of the other atoms, but otherwise its motion is independent or "uncorrelated" with that of the others.

A more realistic wave function is one of the Jastrow-type

$$\psi(\mathbf{x})=\prod_{i=1}^{N}\phi(\mathbf{x}_i-\mathbf{R}_i)\prod_{j<k}f(x_{jk}), \tag{8}$$

where

$$\begin{aligned} f(x_{jk})&\to 0 \quad \text{if} \quad x_{jk}\to 0 \\ f(x_{jk})&\to 1 \quad \text{if} \quad x_{jk}\to\infty . \end{aligned} \tag{9}$$

The factor $f(x_{jk})$ is the "correlation" function. The wave function (8) gives each nucleus a larger effective potential well than (7) since it allows it to move more freely toward and in between neighboring lattice sites. Divergences in the expectation value of $V(\mathbf{x})$ are avoided, even if the $\phi(\mathbf{x}_i-\mathbf{R}_i)$ are overlapping, by the vanishing of $f(x_{jk})$ as $x_{jk}\to 0$.

As normalized single-particle wave functions we choose

$$\phi(\mathbf{x}_i-\mathbf{R}_i)=(2A/\pi)^{3/4}\exp[-A(\mathbf{x}_i-\mathbf{R}_i)^2], \tag{10}$$

where A is a constant.

We write

$$f(x_{jk})=\exp[S(j,k)], \tag{11}$$

and take

$$S(j,k)=-\tfrac{1}{2}(C/x_{jk}{}^n), \tag{12}$$

where C, n are positive constants. It is clear that our $f(x_{jk})$ satisfies the conditions of Eqs. (9). The constants C, A, and n are variational parameters.

III. THE ENERGY EXPECTATION VALUE AND ITS CLUSTER EXPANSION

The expectation value of the energy is given by

$$E=\frac{\int e^S\Phi[-\lambda^2\sum_i\nabla_i^2+V(\mathbf{x})]e^S\Phi d\mathbf{x}}{\int e^{2S}\Phi^2 d\mathbf{x}}, \tag{13}$$

where $\Phi=\prod_i\phi(\mathbf{x}_i-\mathbf{R}_i)$ is real; $S=\sum_{i<j}S(i,j)$, and $d\mathbf{x}=d\mathbf{x}_1d\mathbf{x}_2\cdots d\mathbf{x}_N$. If we integrate by parts we find

$$\int\Phi e^S\nabla_i^2\Phi e^S d\mathbf{x}=\int e^{2S}\{[\tfrac{1}{4}\nabla_i^2(\Phi^2)-(\nabla_i\Phi)^2]+\tfrac{1}{2}\Phi^2\sum_{j(j\neq i)}\nabla_i^2S(i,j)\}d\mathbf{x}. \tag{14}$$

(See Ref. 6 for details. Note the absence of three-particle terms which occur in earlier Jastrow-type calculations.[7]) For our Gaussian wave functions, the term in square brackets of Eq. (14) becomes simply $(-3A\Phi^2)$. By Eq. (12)

$$\nabla_i^2S(i,j)=-(C/2)[n(n-1)/x_{ij}{}^{n+2}]. \tag{15}$$

Thus Eq. (13) becomes

$$E=3A\lambda^2N+\frac{\sum_{i<j}\int e^S\Phi\tilde{V}(i,j)e^S\Phi d\mathbf{x}}{\int e^{2S}\Phi^2 d\mathbf{x}}, \tag{16}$$

where the "effective" two-body potential is

$$\tilde{V}(i,j)=4(x_{ij}{}^{-12}+px_{ij}{}^{-(n+2)}-x_{ij}{}^{-6}), \tag{17}$$

with

$$p=(C/8)\lambda^2n(n-1). \tag{18}$$

We evaluate the matrix elements of $\tilde{V}_{ij}$ in Eq. (16) by means of the cluster expansion technique.[6-8] Following the method of Iwamoto and Yamada[8] we write

$$\frac{\sum_{i<j}\int e^S\Phi\tilde{V}(i,j)e^S\Phi d\mathbf{x}}{\int e^{2S}\Phi^2 d\mathbf{x}}=\frac{\partial}{\partial\beta}\ln I_N(\beta)\Big|_{\beta=0}, \tag{19}$$

where the "generalized normalization integral," $I_N(\beta)$, is defined by

$$I_N(\beta)=\int\exp(2S+\beta\tilde{V})\Phi^2 d\mathbf{x}, \tag{20}$$

and $\tilde{V}=\sum_{i<j}\tilde{V}(i,j)$. Define the "subnormalization" integrals by

$$I_i=(i|i)\equiv\int\phi^2(\mathbf{x}_1-\mathbf{R}_i)d\mathbf{x}_1=1, \tag{21a}$$

$$I_{ij}=(ij|g(1,2)|ij)\equiv\int\int\phi^2(\mathbf{x}_1-\mathbf{R}_i)\times\phi^2(\mathbf{x}_2-\mathbf{R}_j)g(1,2)d\mathbf{x}_1d\mathbf{x}_2, \tag{21b}$$

$$I_{ijk}=(ijk|g(1,2)g(2,3)g(3,1)|ijk), \tag{21c}$$

$$\vdots \qquad\qquad \vdots$$

$$I_{i_1i_2\cdots i_m}=(i_1i_2\cdots i_m|\prod_{1\leqslant\gamma<\delta\leqslant m}g(\gamma,\delta)|i_1i_2\cdots i_m), \tag{21d}$$

where

$$g(\gamma,\delta)=\exp[2S(\gamma,\delta)+\beta\tilde{V}(\gamma,\delta)]. \tag{22}$$

[6] H. W. Jackson and E. Feenberg, Ann. Phys. (N. Y.) 15, 266 (1961).

[7] J. W. Clark and E. Feenberg, Phys. Rev. 113, 388 (1959).

[8] F. Iwamoto and M. Yamada, Progr. Theoret. Phys. (Kyoto) 17, 543 (1957); 18, 345 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Using the subnormalization integrals we define the cluster integrals

$$Y_i=I_i, \tag{23a}$$

$$Y_{ij}=I_{ij}-I_iI_j, \tag{23b}$$

$$Y_{ijk}=I_{ijk}-I_{ij}I_k-I_{jk}I_i-I_{ki}I_j-I_iI_jI_k, \tag{23c}$$

etc.

We find

$$Y_i=1, \tag{24a}$$

$$Y_{ij}=(ij|h(1,2)|ij), \tag{24b}$$

$$Y_{ijk}=(ijk|[h(1,2)h(2,3)h(3,1)+h(1,2)h(3,1)+h(1,2)h(2,3)+h(2,3)h(3,1)]|ijk), \tag{24c}$$

where

$$h(\gamma,\delta)=g(\gamma,\delta)-1. \tag{25}$$

In Ref. 8 it is shown that

$$I_N(\beta)=\prod_i Y_i e^G, \tag{26}$$

where

$$G=\tfrac{1}{2}\sum_{i,j}{}' y_{ij}-\tfrac{1}{4}\sum_{i,j}{}' y_{ij}^2-\tfrac{1}{2}\sum_{i,j,k}{}' y_{ij}y_{jk}+\tfrac{1}{6}\sum_{i,j,k}{}' y_{ijk}+\cdots, \tag{27}$$

and, for example, $y_{ij}=Y_{ij}/Y_iY_j$. The prime on each of the summations of Eq. (27) means omit those terms for which two or more indices are equal. Equation (27) is an expansion in powers of the function $h(\gamma,\delta)$. We will consider only terms with two or fewer factors of $h(\gamma,\delta)$. Those terms symbolized by the dots in Eq. (27) are of higher order in $h(\gamma,\delta)$ than we wish to consider.

We find from Eqs. (16)–(27)

$$E=3A\lambda^2N+\tfrac{1}{2}\sum_{i,j}{}'(ij|W(1,2)|ij)-\tfrac{1}{2}\sum_{i,j}{}'(ij|W(1,2)|ij)(ij|\eta(1,2)|ij)-\sum_{i,j,k}{}'(ij|\eta(1,2)|ij)(jk|W(1,2)|jk)+\sum_{i,j,k}{}'(ijk|W(1,2)\eta(1,3)|ijk)+\cdots, \tag{28}$$

where

$$W(\gamma,\delta)=e^{2S(\gamma,\delta)}\bar{V}(\gamma,\delta), \tag{29}$$

and

$$\eta(\gamma,\delta)=h(\gamma,\delta)|_{\beta=0}=(e^{2S(\gamma,\delta)}-1). \tag{30}$$

Equation (28) includes terms no higher than first order in the function $\eta(\gamma,\delta)$. Succeeding terms will be higher order in $\eta(\gamma,\delta)$. We will discuss the contribution of such terms later. Note that the new "effective" potential $W(\gamma,\delta)$ contains a factor

$$\exp[2S(\gamma,\delta)]=\exp[-Cx_{\gamma\delta}{}^{-n}],$$

which insures convergent results for the matrix elements in Eq. (28).

IV. REVIEW OF PREVIOUS THEORETICAL TREATMENTS

Bernardes has made two variational calculations, each of which involves the use of a trial wave function which is a product of single-particle wave functions each spherically symmetric around a lattice site. In his first calculation[1] he found an energy per particle of -420 cal/mole compared to the experimental value of -450 ± 10 cal/mole.[9] His second calculation[2] gave a value of -445 cal/mole in good agreement with experiment. Nosanow and Shaw,[2] however, have done a Hartree calculation resulting in a theoretical value of -431 cal/mole. They used the same potential parameters, σ and ϵ, in Eq. (4), as did Bernardes. The Hartree technique used by Nosanow and Shaw yields the best result in a variational computation of the energy with a trial wave function which is a product of single-particle wave functions each of which is spherically symmetric about a lattice site. We may speculate that Bernardes' first result does not agree with that of Nosanow and Shaw either because his wave function lacks sufficient variational flexibility or because an expansion he makes does not converge sufficiently rapidly. As he discounts both of these possibilities we are unable to resolve the discrepancy. Bernardes' second calculation leads to an energy which is actually lower than the result of the Hartree technique. Since our results, given in the next section, agree with those of the Hartree technique we are inclined to believe that Bernardes' agreement with experiment is fictitious. Possibly he did not keep a sufficient number of terms in an expansion made in powers of λ.

V. NUMERICAL TECHNIQUES AND RESULTS

As a first approximation we keep only the first two terms of Eq. (28), giving

$$E=3A\lambda^2N+\sum_{i<j}(ij|W(1,2)|ij). \tag{31}$$

We have

$$(ij|W(1,2)|ij)=\int\int\phi^2(\mathbf{x}_1-\mathbf{R}_i)\phi^2(\mathbf{x}_2-\mathbf{R}_j)\times W(x_{12})d\mathbf{x}_1d\mathbf{x}_2$$

$$=\left(\frac{2A}{\pi}\right)^3\int\int\exp[-2A(\mathbf{x}_1-\mathbf{R}_i)^2]\times\exp[-2A(\mathbf{x}_2-\mathbf{R}_j)^2]\times W(x_{12})d\mathbf{x}_1d\mathbf{x}_2. \tag{32}$$

[9] E. R. Dobbs and G. O. Jones, *Reports on Progress in Physics* (The Physical Society, London, 1957), Vol. 20, p. 516.

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

Let $\mathbf{z}_1=\mathbf{x}_1-\mathbf{x}_2$, $\mathbf{z}_2=\mathbf{x}_2-\mathbf{R}_i$. Then

$$(ij|W(1,2)|ij)=\left(\frac{2A}{\pi}\right)^3\int d\mathbf{z}_1 W(|\mathbf{z}_1|)$$
$$\times\exp[-2A\mathbf{z}_1^2]\int d\mathbf{z}_2\exp[-4A(\mathbf{z}_2+\mathbf{R}_{ij})^2]$$
$$\times\exp[-2A(\mathbf{z}_2^2+2\mathbf{z}_1\cdot\mathbf{z}_2)], \quad (33)$$

with $\mathbf{R}_{ij}=\mathbf{R}_i-\mathbf{R}_j$. Carrying out the $\mathbf{z}_2$ integration yields

$$(ij|W(1,2)|ij)=(2A/\pi)^3\int d\mathbf{z}_1 W(|\mathbf{z}_1|)$$
$$\times\exp[-A(\mathbf{z}_1+\mathbf{R}_{ij})^2]$$
$$=(A/\pi)^{1/2}(1/R_{ij})\int_0^\infty dz\, zW(z)$$
$$\times\{\exp[-A(z-R_{ij})^2]$$
$$-\exp[-A(z+R_{ij})^2]\}, \quad (34)$$

where $R_{ij}=|\mathbf{R}_{ij}|$ and $z=|\mathbf{z}_1|$.

Our procedure is to choose initial values for n and C and minimize E as a function of A. We then choose a new value for C and again minimize with respect to A. In this way we find the minimum of E as a function of C. The whole procedure is repeated for a second value of n, etc. All of the above is done for various values of the nearest-neighbor distance R until we find the equilibrium or zero pressure nearest-neighbor distance R_0 which is the value of R when E is an absolute minimum. The procedure is greatly simplified by the fact that the energy expectation value has a broad minimum as a function of C and n. The computations were performed numerically using an IBM-7090 computer.

In computing the lattice sums indicated in Eq. (31) we found sufficient accuracy was obtained if we summed over the first sixty "shells" of particles, all particles on a "shell" being a given distance from the particle at the origin. This involved summing over the nearest 3126 neighbors to each nucleus. Values for the numbers of particles on the shells and for the shell distances for the face-centered cubic lattice were taken from Ref. 10.

First we make use of Bernardes' potential constants $\sigma=2.74$ Å and $\epsilon=50.0\times10^{-16}$ erg. (Second virial coefficient gas-phase data give $\sigma=2.75$ Å, $\epsilon=49.2\times10^{-16}$ erg.[11]) Initially we take the convenient value $n=10$. We find that the value of C which minimizes the energy at a nearest-neighbor distance near the experimental equilibrium value is less than 10^{-4}. In fact, we find that the lowest energy value given by Eq. (31) occurs if, in our numerical computations, we set $C=0$ and begin our integration in Eq. (34) at that value of z for which $\exp[-A(z-R_{ij})^2]$ is negligibly small, say, less than 10^{-5} in comparison with its maximum of unity. Of course, C cannot be precisely zero since then the matrix element $(ij|W(1,2)|ij)$ would diverge. We will describe this situation by saying $C=0+$.

A good measure of the importance of correlation is the relative size of the term $px_{ij}^{-(n+2)}$ in comparison with the other potential terms in Eq. (17). For $n=10$, the ratio of this term to x_{ij}^{-12} is $p=90(\lambda^2C/8)<10^{-5}$. Because C is so small the result does not depend on the value of n chosen. To demonstrate this statement explicitly we consider the case $n=4$. Again we find $C=10^{-4}$ and the ratio of $px_{ij}^{-(n+2)}$ to x_{ij}^{-6} is less than 10^{-6}. Thus we have the important result that near equilibrium correlation effects are negligible in solid neon. Table I summarizes the results.

Because the value of C minimizing the approximate E value of Eq. (31) is small, the neglected terms in the expansion, Eq. (28), do not contribute. To see this consider the matrix element $(ij|\eta(1,2)|ij)$. Using the derivation leading to Eq. (34), we can show

$$(ij|\eta(1,2)|ij)=(A/\pi)^{1/2}(1/R_{ij})$$
$$\times\int_0^\infty dz\, z\{\exp[-Cz^{-n}]-1\}\{\exp[-A(z-R_{ij})^2]$$
$$-\exp[-A(z+R_{ij})^2]\}. \quad (35)$$

The factor $\exp[-Cz^{-(n+2)}]$ is unity over the entire range of integration for which the Gaussian functions are appreciable, so that $(ij|\eta(1,2)|ij)$ is essentially zero. Similar arguments hold for the other terms first order in $\eta(\gamma,\delta)$ in Eq. (28). Other terms in the expansion are of even higher order in $\eta(\gamma,\delta)$ and also do not contribute.

TABLE I. Energy expectation value as a function of the variational parameters C and n, using Bernardes' potential constants[a] $\sigma=2.74$Å, $\epsilon=50.0\times10^{-16}$ erg. All quantities are expressed in reduced units, length in units of σ, energy in units of ϵ. A_0 is the value of the parameter A which yields the minimum energy per particle, E_0/N, for a given C and n, at a fixed value $R=1.138$ of the nearest-neighbor distance near equilibrium. (For the sake of comparison we have quoted values of E_0/N to four significant figures. Although we claim only three figure absolute accuracy, the relative accuracy between the values of E_0/N for various values of C is four figures.)

n	C	A_0	$-E_0/N$
10	0+[b]	94.0	5.984
	10^{-4}	94.0	5.984
	10^{-3}	94.0	5.982
	10^{-2}	93.6	5.968
4	0+[b]	94.0	5.984
	10^{-4}	94.0	5.983
	10^{-3}	94.0	5.980
	10^{-2}	93.7	5.953

[a] Reference 1.
[b] Explained in Sec. V of text.

[10] J. Hirschfelder, C. Curtis, and R. Bird, *Molecular Theory of Gases and Liquids* (John Wiley & Sons Inc., New York, 1954), p. 1037.
[11] J. de Boer, Physica 14, 139 (1948).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Comparison of various theoretical results with experiment for the energy per particle $\epsilon E_0/N$ and the equilibrium nearest-neighbor distance σR_0.

	Theory $\sigma=2.74$ Å, $\epsilon=50.0\times10^{-16}$ erg			Theory $\sigma=2.75$ Å, $\epsilon=51.7\times10^{-16}$ erg		Experiment
	Bernardes (Ref. 1)	Bernardes (Ref. 3)	Nosanow-Shaw (Ref. 2)	This work	This work	Refs. 9 and 12
$-\epsilon E_0/N$ (cal/mole)	420	445[a]	431[b]	431	449	450±10
σR_0 (Å)	3.14	3.13[a]	—	3.12	3.13	3.13

[a] See Sec. IV of text.
[b] Computed at the experimental value of R_0 only.

If we transform from our reduced units we find the minimum energy per particle at the equilibrium density to be $\epsilon E_0/N=-431$ cal/mole, in precise agreement with the Hartree computation. This result shows that we do not need additional variational parameters in our wave function. The nearest neighbor distance at equilibrium is found to be $\sigma R_0=3.12$ Å in comparison with the experimental values[12] of 3.13 Å.

To find the pressure-volume relationship for the solid, we use $P=-\partial E/\partial V$, where P is the pressure, and V the volume. We analytically differentiate Eq. (31) with respect to the nearest-neighbor distance R and use

$$P=-\frac{\partial E}{\partial V}=-\frac{2^{1/2}}{3}\frac{1}{R^2}\frac{\partial E}{\partial R} \tag{36}$$

which holds for the face-centered cubic lattice. In this equation P is expressed in units of ϵ/σ^3. We find

$$P=(8A/\pi)^{1/2}(1/3R^3)(F_0-2AF_1)\,, \tag{37}$$

where

$$F_0=\sum_{i<j}(1/R_{ij})\int_0^\infty dz\, zW(z)\{\exp[-A(z-R_{ij})^2]-\exp[-A(z+R_{ij})^2]\}\,, \tag{38}$$

$$F_1=\sum_{i<j}\int_0^\infty dz\, zW(z)\{(z-R_{ij})\exp[-A(z-R_{ij})^2]-(z+R_{ij})\exp[-A(z+R_{ij})^2]\}\,. \tag{39}$$

Given an R value we find that value of A, which minimizes E (with $C=0+$ and $n=10$) and compute the corresponding pressure by Eq. (37). (Note that by taking $C=0+$, we have assumed correlation to be negligible at all pressures.) Our results for Bernardes' σ and ϵ are shown as curve A in Fig. 1, where we have plotted pressure in kg/cm² versus V/V_0, V_0 being the molar volume at equilibrium. The dots are the experimental data of Stewart.[13] We note that there is some disagreement between curve A and the experimental points in contrast to the good agreement each of Bernardes' calculations give. (Bernardes' curves are not shown in Fig. 1.)

The question arises concerning the possibility of obtaining agreement between our theory and all of the experimental data by merely adjusting the potential constants σ and ϵ by a few percent. Consider the values $\sigma=2.75$ Å and $\epsilon=51.7\times10^{-16}$ ergs. The performance of the same operations as described above gives a value of -449 cal/mole for the energy per particle at an equilibrium distance of 3.13 Å. Table II compares the various theoretical results with experiment. Curve B in Fig. 1 is the resulting P-V relationship which is in good agreement with the data up to about 15×10^3 kg/cm². Above that pressure it is possible that correlation effects become nonnegligible. Our discussion of correlation applies only to equilibrium.

VI. DISCUSSION

Our results show that it is possible to explain the properties of solid neon on the basis of a two-body (Mie–Lennard-Jones 12-6) potential and a simple product trial wave function. Correlation effects are

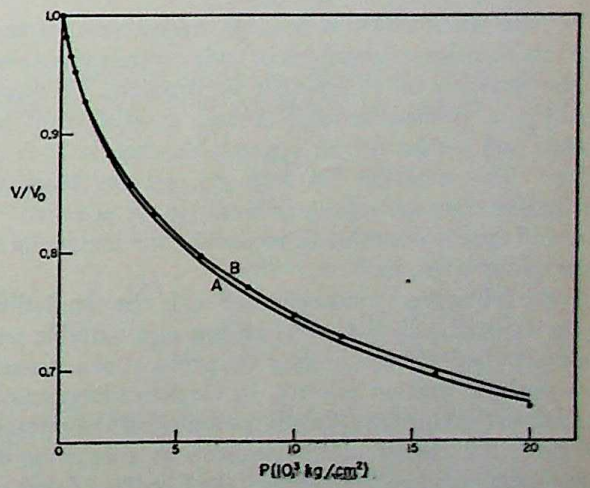


FIG. 1. Pressure-volume relationship for solid neon. The dots are the experimental values of Stewart (Ref. 13). Curves A and B are the theoretical relationships corresponding, respectively, to the potential constants of Bernardes, $\sigma=2.74$ Å, $\epsilon=50.0\times10^{-16}$ erg, and to those discussed in the text, $\sigma=2.75$ Å, $\epsilon=51.7\times10^{-16}$ erg.

[12] D. G. Henshaw, Phys. Rev. 111, 1470 (1958).
[13] J. W. Stewart, Phys. Chem. Solids 1, 146 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

negligible at low pressures. The values of the potential constants which give a fit to the data, $\sigma=2.75$ Å, $\epsilon=51.7\times10^{-16}$ ergs, differ at most by only a few percent from those given by gas phase data, $\sigma=2.75$ Å, $\epsilon=49.2\times10^{-16}$ ergs.

The fact that different constants are needed to explain the gas and solid-phase data of neon should not be at all surprising. The Mie–Lennard-Jones potential is a phenomenological interaction and should be expected to vary somewhat from one type of phenomenon to another.[14] A similar situation occurs, for example, in the treatment of argon. Second virial coefficient data[15] give, $\epsilon=165\times10^{-16}$ ergs, gas viscosity data[16] give $\epsilon=171\times10^{-16}$ ergs, and crystal data[17] give $\epsilon=169\times10^{-16}$ ergs.[18]

[14] See, for example, Ref. 10, p. 208.
[15] E. Whalley and W. G. Schneider, J. Chem. Phys. 23, 1644 (1955).
[16] H. L. Johnston and E. R. Grilly, J. Phys. Chem. 46, 938 (1942).

ACKNOWLEDGMENTS

We would like to thank the Digital Computation Division (ASNCD) of the Aeronautical Systems Division at Wright-Patterson Air Force Base for their assistance in processing the computer programs. We are grateful to Professor Eugene Feenberg for his encouragement and for a critical reading of the manuscript.

[17] C. Domb and I. J. Zucker, Nature 178, 484 (1956).
[18] A small change of potential parameters may not account for the experimental data of the heavier inert gas solids. See the P-V data and discussion of C. A. Swenson for xenon [*International Conference on the Physics and Chemistry of High Pressures* (Society of the Chemical Industry, London, June, 1962)]. It may be necessary to include the effects of many-body forces to correlate the properties of solid xenon.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Positronium in a LiH Crystal*

S. M. Neamtan and R. I. Verrall
Department of Mathematical Physics, The University of Manitoba, Winnipeg, Canada
(Received 9 December 1963; revised manuscript received 30 January 1964)

The question as to whether the existence of a bound positronium-like system in an ionic crystal is energetically possible is considered with particular reference to lithium hydride. The crystal field of LiH is simulated by a potential expressed as a sum of central field potentials centered on the ion sites. An upper bound is obtained for the ground-state energy of a system of an electron and a positron in Coulomb interaction with each other and moving in the crystal potential, and comparison is made with estimates of the energies of the positron and the electron moving independently in the crystal. The result is energetically favorable to the formation and persistence of positronium in the crystal through capture by a positron of an electron from the valence band.

I. INTRODUCTION

CONSIDERATION of possible mechanisms for the annihilation of positrons in ionic crystals gives rise to the question as to whether positronium, or more precisely a positronium-like bound system, can be formed and persist for an appreciable time in such a crystal. This question has formerly, on the basis of qualitative arguments, been answered in the negative.[1,2] It would appear desirable to reconsider the matter on a more quantitative basis.

In the following, we consider in Sec. II the simulation of the crystal field of LiH by means of a suitable potential. In Sec. III we consider the problem of a system of a positron and an electron in Coulomb interaction moving in the constant periodic potential of the crystal, the potential energy of the positron as a function of position being taken to be simply the negative of that of the electron. By treating the effect of the crystal field as a perturbation of the Hamiltonian for free positronium, the energy shift of the ground state is calculated to second order of perturbation theory. An improved result is then obtained in the form of an upper bound to the ground-state energy by calculating the average of the Hamiltonian with the aid of a wave function given by perturbation theory. In Sec. IV the procedure used in estimating the ground-state energy of a positron in the crystal field is outlined. In Sec. V use is made of the results obtained in a discussion of the energetics of positronium formation in the crystal. Section VI contains an assessment of the results and their relation to experiment.

II. CRYSTAL POTENTIAL OF LiH

The lithium hydride crystal, being an ionic crystal, has a relatively open structure. We therefore make the approximation of treating the potential as being made up of a sum of spherically symmetric potentials centered on the ion sites. The first problem, then, is to obtain an

* Supported by the National Research Council of Canada.
[1] R. A. Ferrell, Rev. Mod. Phys. 28, 308 (1956).
[2] P. R. Wallace, *Solid State Physics* (Academic Press Inc., New York, 1960), Vol. 10.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

approximation to the potential due to a lithium ion and that due to a hydride ion.

We shall take the potential due to a lithium ion to be of the form

$$\mathcal{U}_1(R)=(e/R)+(2e/R)e^{-\alpha R}, \quad (1)$$

where e is the charge on a proton and α is a parameter which is to be determined. This potential has the correct behavior for both large and small R. It has the advantage of being simple in its mathematical form. Similarly, the potential of the hydride ion is taken to be

$$\mathcal{U}_2(R)=(2e/R)e^{-\beta R}-(e/R), \quad (2)$$

where β is a parameter which is to be determined.

The parameter α is obtained by observing that the potential $\mathcal{U}_1(R)$ is effectively the potential in which the third electron of a lithium atom moves. The ground-state energy of this third electron is known. We may, therefore, set up the Schroedinger equation for this particle, wherein only the parameter α is unknown. With a fixed value of α we may solve the equation numerically. The parameter α can be varied until we obtain a solution which has one radial node and which behaves properly as $r\rightarrow\infty$. This procedure was carried out on a computer, and the value for α obtained was

$$\alpha=4.152\ \text{Å}^{-1}. \quad (3)$$

The value for the parameter β in the hydride ion potential was obtained in a different fashion. Using a wave function for the hydride ion, the charge density is

$$\rho(\mathbf{R}_1)=-2e\int_\tau|\psi(\mathbf{R}_1,\mathbf{R}_2)|^2 d\tau_2. \quad (4)$$

The potential due to this charge density, $v(R)$, is given by

$$\nabla^2 v(\mathbf{R})=-4\pi\rho(\mathbf{R}). \quad (5)$$

The wave function for the hydride ion was taken to be of the form used by G. Darewych.[3]

$$\psi(\mathbf{R}_1,\mathbf{R}_2)=(1+\nu R_{12})(e^{-(\lambda R_1+\mu R_2)}+e^{-(\mu R_1+\lambda R_2)}), \quad (6)$$

where

$$R_{12}=|\mathbf{R}_2-\mathbf{R}_1|, \quad (7)$$

and

$$\lambda=0.9033\ \text{Å}^{-1},\quad \mu=2.0315\ \text{Å}^{-1},\quad \nu=0.5898\ \text{Å}^{-1}. \quad (8)$$

Using this wave function, it is possible to carry out the integration in Eq. (4) and to solve the Poisson equation to give an explicit form of the potential. This potential was approximated with high accuracy by the form

$$\mathcal{U}_2(R)=(2e/R)e^{-\beta R}-(e/R). \quad (9)$$

[3] G. Darewych (private communication).

The value of β for the best fit is

$$\beta=1.715\ \text{Å}^{-1}. \quad (10)$$

The crystal potential is now taken to be

$$V(\mathbf{R})=\sum_j[\mathcal{U}_1(\mathbf{R}-\mathbf{R}_j)+\mathcal{U}_2(\mathbf{R}-\mathbf{R}_j-\mathbf{a})], \quad (11)$$

where $\mathbf{R}_j$ is the position vector of the jth positive ion and $\mathbf{a}$ is the position vector of a negative ion with respect to an origin at the location of a positive ion.

III. GROUND-STATE ENERGY OF POSITRONIUM IN THE CRYSTAL

We shall now investigate the problem of positronium in this potential. The Hamiltonian for the two-particle system in this field is

$$H=-(\hbar^2/2m)(\nabla_1^2+\nabla_2^2)-(e^2/|\mathbf{r}_2-\mathbf{r}_1|)+eV(\mathbf{r}_1)-eV(\mathbf{r}_2), \quad (12)$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the positron and electron, respectively. It is to be noted that the potential energies of the electron and of the positron due to the crystal field differ only in sign.

The Hamiltonian may be expressed in terms of the center-of-mass coordinate $\mathbf{R}$ and the relative coordinate $\mathbf{r}$, where we have

$$\mathbf{R}=(\mathbf{r}_1+\mathbf{r}_2)/2 \quad (13)$$

and

$$\mathbf{r}=\mathbf{r}_2-\mathbf{r}_1. \quad (14)$$

The Hamiltonian becomes

$$H=-\frac{\hbar^2}{4m}\nabla_R^2-\frac{\hbar^2}{m}\nabla_r^2-\frac{e^2}{r}+eV\left(\mathbf{R}-\frac{\mathbf{r}}{2}\right)-eV\left(\mathbf{R}+\frac{\mathbf{r}}{2}\right). \quad (15)$$

In accordance with perturbation theory, we write this Hamiltonian as the sum of two parts:

$$H=H_0+H_1, \quad (16)$$

where

$$H_0=-\frac{\hbar^2}{4m}\nabla_R^2-\frac{\hbar^2}{m}\nabla_r^2-\frac{e^2}{r} \quad (17)$$

and

$$H_1=eV\left(\mathbf{R}-\frac{\mathbf{r}}{2}\right)-eV\left(\mathbf{R}+\frac{\mathbf{r}}{2}\right). \quad (18)$$

H_1 will be treated as a perturbation on the Hamiltonian H_0.

H_0 consists of the sum of two noninteracting parts,

$$-(\hbar^2/4m)\nabla_R^2 \quad (19)$$

and

$$-(\hbar^2/m)\nabla_r^2-(e^2/r). \quad (20)$$

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

The first is the Hamiltonian of a free particle and the second is the Hamiltonian of the positronium atom. Therefore, the eigenfunctions of H_0 may be written

$$\psi_{\mathbf{k},nlm}=e^{i\mathbf{k}\cdot\mathbf{R}}\phi_{nlm}(\mathbf{r}), \tag{21}$$

where $\phi_{nlm}(\mathbf{r})$ is a positronium wave function, where the quantum numbers n, l, m have the usual meaning. The energy corresponding to this state is

$$(\hbar^2k^2/4m)+E_n, \tag{22}$$

where E_n is the energy of the nth level of positronium.

Let us examine the effect of the perturbation on the ground-state energy. The first-order perturbation energy is

$$[\psi_{0,100}(\mathbf{r}),H_1\psi_{0,100}(\mathbf{r})]. \tag{23}$$

Since H_1 is of odd parity in the $\mathbf{r}$ space, this vanishes. From similar parity arguments, it can be seen that all odd-order perturbation energies vanish. Thus the second-order perturbation, which we shall now calculate, is in fact accurate to third order.

The second-order perturbation energy is

$$\sum_{\mathbf{k},nlm}\frac{|[\psi_{0,100}(\mathbf{r}),H_1\psi_{\mathbf{k},nlm}(\mathbf{r})]|^2}{E_1-[(\hbar^2k^2/4m)+E_n]}. \tag{24}$$

The inner product in the above expression involves an integration over both the $\mathbf{R}$ and the $\mathbf{r}$ spaces. The functions of $\mathbf{R}$ are not quadratically integrable, so we must resort to a box quantization procedure. The inner product becomes

$$\operatorname*{Lim}_{L\to\infty}\frac{1}{L^3}\int_{\tau_R,\tau_r}\phi_{100}^*(r)\left[eV\left(\mathbf{R}-\frac{\mathbf{r}}{2}\right)-eV\left(\mathbf{R}+\frac{\mathbf{r}}{2}\right)\right]\times\phi_{nlm}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{R}}d\tau_R d\tau_r, \tag{25}$$

where $1/L^3$ is the square of the normalization constant used in the box quantization procedure. By suitable shifts of origin, expression (25) can be reduced to the form

$$\left[\operatorname*{Lim}_{L\to\infty}\frac{1}{L^3}\int_{\tau_R}e^{i\mathbf{k}\cdot\mathbf{R}}eV(\mathbf{R})d\tau_R\right]\times\left[\int_{\tau_r}\phi_{100}^*(r)\phi_{nlm}(\mathbf{r})(e^{i\mathbf{k}\cdot(\mathbf{r}/2)}-e^{-i\mathbf{k}\cdot(\mathbf{r}/2)})d\tau_r\right]. \tag{26}$$

Now $V(\mathbf{R})$ has the periodicity of the crystal. It may therefore be expanded as a Fourier series.

$$V(\mathbf{R})=\sum C_{\mathbf{K}}e^{2\pi i\mathbf{K}\cdot\mathbf{R}}, \tag{27}$$

where $\mathbf{K}$ is a vector of the reciprocal lattice.

The integral

$$\operatorname*{Lim}_{L\to\infty}\frac{1}{L^3}\int_{\tau_R}e^{i\mathbf{k}\cdot\mathbf{R}}V(\mathbf{R})d\tau_R \tag{28}$$

is equal to $C_{-\mathbf{K}}$, if $\mathbf{k}$ is 2π times a reciprocal lattice vector and is zero otherwise. Also, it is obvious that $C_{\mathbf{K}}=C_{-\mathbf{K}}$. Substituting these results into Eq. (26), we find that the second-order perturbation, expression (24), reduces to

$$4\sum_{\mathbf{K}}|C_{\mathbf{K}}|^2\sum_{nlm}\frac{\left|\int\phi_{100}^*(\mathbf{r})\phi_{nlm}(\mathbf{r})\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}}{2}\right)d\tau_r\right|^2}{E_1-\left(\frac{\hbar^2k^2}{4-m}+E_n\right)}. \tag{29}$$

Now $E_n=E_1/n^2$. As n increases E_n quickly becomes small. We set E_n equal to zero, in the above, in order to facilitate summations. We shall come back later and correct for the first few terms. Using this approximation, (29) may now be written

$$4\sum_{\mathbf{K}}|C_{\mathbf{K}}|^2\sum_{nlm}\frac{\int\phi_{100}^*(r)\phi_{100}(r')\phi_{nlm}(\mathbf{r})\phi_{nlm}^*(\mathbf{r}')\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}}{2}\right)\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}'}{2}\right)d\tau_r d\tau_{r'}}{E_1-\frac{\hbar^2k^2}{4m}}. \tag{30}$$

We now sum over n, l, and m. Since $\{\phi_{nlm}(\mathbf{r})\}$ is a complete orthonormal set of functions, we have

$$\sum_{nlm}\phi_{nlm}^*(\mathbf{r}')\phi_{nlm}(\mathbf{r})=\delta(\mathbf{r}'-\mathbf{r}). \tag{31}$$

Integration over the $\mathbf{r}'$ space reduces the integral to

$$4\sum_{\mathbf{K}}|C_{\mathbf{K}}|^2\frac{\int|\phi_{100}(r)|^2\sin^2\left(\frac{\mathbf{k}\cdot\mathbf{r}}{2}\right)d\tau_r}{E_1-\frac{\hbar^2k^2}{4m}}, \tag{32}$$

which is equal to

$$2\sum_{\mathbf{K}}|C_{\mathbf{K}}|^2\frac{\{1-[1/(1+a_0^2k^2)^2]\}}{E_1-(\hbar^2k^2/4-m)}. \tag{33}$$

We must now evaluate $C_{\mathbf{K}}$. From Eq. (27) we obtain

$$C_{\mathbf{K}}=\frac{1}{\Delta}\int V(\mathbf{R})e^{-2\pi i\mathbf{K}\cdot\mathbf{R}}d\tau_R, \tag{34}$$

where Δ is the volume of a unit cell of the crystal lattice and the integration is carried out over this volume.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Substituting Eq. (11) in (34) we obtain

$$C_{\mathbf{K}}=\frac{1}{\Delta}\int\sum_j[\mathcal{V}_1(\mathbf{R}-\mathbf{R}_j)+\mathcal{V}_2(\mathbf{R}-\mathbf{R}_j-\mathbf{a})]\times e^{-2\pi i\mathbf{K}\cdot\mathbf{R}}d\tau_R. \quad (35)$$

It is easily seen that this is equal to

$$\frac{1}{\Delta}\int[\mathcal{V}_1(\mathbf{R})+\mathcal{V}_2(\mathbf{R}-\mathbf{a})]e^{-2\pi i\mathbf{K}\cdot\mathbf{R}}d\tau_R, \quad (36)$$

where this integration is carried out over all space. We now let

$$\mathbf{K}=n_1\mathbf{b}_1+n_2\mathbf{b}_2+n_3\mathbf{b}_3, \quad (37)$$

where $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ are a set of basis vectors in the reciprocal lattice, and n_1, n_2, and n_3 are integers. On substituting the functions $V_1(R)$ and $V_2(R)$ into the integral we obtain

$$C_{\mathbf{K}}=\frac{2\pi e^2}{L^3}\left\{\frac{2}{k^2+\alpha^2}+\frac{1}{k^2}+(-1)^{n_1+n_2+n_3}\left[\frac{2}{k^2+\beta^2}-\frac{1}{k^2}\right]\right\}, \quad (38)$$

where L is the shortest distance between positive and negative ions in the crystal. It is to be remembered that

$$\mathbf{k}=2\pi\mathbf{K}=2\pi(n_1\mathbf{b}_1+n_2\mathbf{b}_2+n_3\mathbf{b}_3). \quad (39)$$

Substitution of this expression for $C_{\mathbf{K}}$ into Eq. (33) leaves only the summation over $\mathbf{K}$ to be performed.

It is recalled that this value is an approximation to expression (29) wherein we took E_n to be zero. We shall now go back and correct for the first few values of n.

From parity considerations it is clear that the integral in expression (29) vanishes when l is even. We therefore need consider only the odd values of l in the correction calculations. This implies immediately that we need not consider $n=1$ for which l has only the value 0. The correction term to be added to (33) is, therefore,

$$4\sum_{\mathbf{K}}|C_{\mathbf{K}}|^2\sum_{n=2}^{5}\sum_{\text{odd }l,m}\int\phi_{100}^*(r)\phi_{100}(r')\phi_{nlm}(\mathbf{r})\phi_{nlm}^*(\mathbf{r}')\times\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}}{2}\right)\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}'}{2}\right)d\tau_{r'}d\tau_r\times\left\{\frac{1}{E_1-E_n-\dfrac{\hbar^2k^2}{4m}}-\frac{1}{E_1-\dfrac{\hbar^2k^2}{4m}}\right\}, \quad (40)$$

where the summation over n is terminated at $n=5$.

Consider the sum

$$\sum_{\text{odd }l,m}\int\phi_{100}^*(r)\phi_{100}(r')\phi_{nlm}(\mathbf{r})\phi_{nlm}^*(\mathbf{r}')\times\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}}{2}\right)\sin\left(\frac{\mathbf{k}\cdot\mathbf{r}'}{2}\right)d\tau_r d\tau_{r'}. \quad (41)$$

We express $\sin(\mathbf{k}\cdot\mathbf{r}/2)$ in terms of complex exponentials and use the relation

$$e^{i\mathbf{k}\cdot\mathbf{r}}=\sum_{l'=0}^{\infty}(2l'+1)i^{l'}P_{l'}(\cos\theta)j_{l'}(kr), \quad (42)$$

where $j_{l'}(kr)$ is a spherical Bessel function. We also write $\phi_{nlm}(\mathbf{r})$ as

$$\phi_{nlm}(r)=R_{nl}(r)Y_l^m(\theta,\phi). \quad (43)$$

On summation over m and integration over angle, the expression (41) reduces to

$$4\pi\sum_{\text{odd }l}(2l+1)\left[\int\phi_{100}(r)R_{nl}(r)j_l\left(\frac{kr}{2}\right)r^2dr\right]^2. \quad (44)$$

The energy shift given by Eq. (40) therefore reduces to

$$16\pi\sum_{\mathbf{K}}\sum_{n=2}^{5}|C_{\mathbf{K}}|^2\sum_{\text{odd }l}(2l+1)\times\left[\int\phi_{100}(r)R_{nl}(r)j_l\left(\frac{kr}{2}\right)r^2dr\right]^2\times\left[\frac{1}{E_1-E_n-(\hbar^2k^2/4m)}-\frac{1}{E-(\hbar^2k^2/4m)}\right], \quad (45)$$

wherein the integrals can be explicitly evaluated. The summations in expressions (33) and (45) were evaluated on a computer.

These computations yield a second-order energy shift of -4.51 eV. Since the unperturbed energy is -6.802 eV, this places the energy of the positronium system at $-6.802-4.51=-11.31$ eV.

Since this energy shift is quite large, it was decided to obtain a check on the result by calculating an upper bound to the energy. This was done by evaluating the expected average of the Hamiltonian using the wave function given by the perturbation theory. This perturbed ground-state wave function is given by

$$\psi=\psi_{0,100}+\sum_{klm}\frac{(\psi_{0,100},H_1\psi_{k,nlm})}{E_1-E_n-(\hbar^2k^2/4m)}\psi_{k,nlm}. \quad (46)$$

Using this wave function, we wish to calculate $(\psi,H\psi)/$

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

(ψ,ψ). The numerator reduces to

$$(\psi,H\psi)=E_1+2\sum_{\mathbf{k},nlm}\frac{|(\psi_{0,100},H_1\psi_{\mathbf{k},nlm})|^2}{\left[E_1-E_n-\dfrac{\hbar^2k^2}{4m}\right]}$$

$$=\sum_{\mathbf{k},nlm}\frac{[E_n+(\hbar^2k^2/4m)]|(\psi_{0,100},H_1\psi_{\mathbf{k},nlm})|^2}{[E_1-E_n-(\hbar^2k^2/4m)]^2}. \quad (47)$$

The denominator reduces to

$$(\psi,\psi)=1+\sum_{\mathbf{k},nlm}\frac{|(\psi_{0,100},H_1\psi_{\mathbf{k},nlm})|^2}{[E_1-E_n-(\hbar^2k^2/4m)]^2}. \quad (48)$$

The first sum in Eq. (47) is the same as that in expression (24). The other two sums in Eqs. (47) and (48) are evaluated in much the same fashion. The result of this calculation is

$$(\psi,H\psi)/(\psi,\psi)=-13.43\text{ eV}. \quad (49)$$

It is to be recalled that this is an upper bound to the energy.

It can be shown that the error in the above result, arising from the approximation procedure used in the summations, is less than 0.01 eV. If, in (47) and (48) we attempt to approximate by taking the discrete sum only up to $n=5$, the result obtained for $(\psi,H\psi)/(\psi,\psi)$ is -7.044 eV. The large discrepancy between this and the more accurate value obtained above indicates that there is a substantial contribution to the wave function ψ from the higher excited states of positronium. This, in turn, implies (since the $l=0$ states do not contribute) that there is a significant departure of the wave function from spherical symmetry.

IV. GROUND-STATE ENERGY OF A POSITRON IN THE CRYSTAL

We must now estimate the ground-state energy of a positron in the simulated potential. An upper bound to this energy is calculated by making use of the variational principle. Since we are seeking the ground state, the wave function must have all the symmetries of the potential. A trial wave function ψ is expanded into a Fourier series, where the basis functions have the form

$$\sum_{P(n_1,n_2,n_3)}P(n_1,n_2,n_3)\cos\frac{2\pi n_1X}{a}\times\cos\frac{2\pi n_2Y}{a}\cos\frac{2\pi n_3Z}{a}. \quad (50)$$

The X, Y, Z coordinates are chosen along the fundamental lattice vectors, each of length a. The symbol $P(n_1,n_2,n_3)$ represents a permutation of the three integers n_1, n_2, and n_3. The sum is carried out over all six permutations. This function has the complete symmetry of the potential.

The Fourier coefficients in the expansion of ψ are used as the variable parameters in the variational technique.

The Hamiltonian for the positron is

$$H=-(\hbar^2/2m)\nabla^2+eV(\mathbf{R}), \quad (51)$$

where $V(\mathbf{R})$ is the crystal potential obtained previously. It can be easily seen that the inner product $(\psi,H\psi)$ is a homogeneous quadratic function of the Fourier coefficients, i.e.,

$$(\psi,H\psi)=\sum_{ij}a_{ij}C_iC_j, \quad (52)$$

where C_i is a Fourier coefficient and a_{ij} is a constant. Similarly, the inner product (ψ,ψ) has the form

$$(\psi,\psi)=\sum_i b_i{}^2C_i{}^2, \quad (53)$$

where the coefficients $b_i{}^2$ are constants.

The variational procedure involves minimizing $(\psi,H\psi)$ with respect to the variable parameters. In other words, we wish to minimize $(\psi,H\psi)$ subject to the condition that $(\psi,\psi)=1$. By the introduction of a Lagrange undetermined multiplier this problem reduces to finding the smallest eigenvalue of the matrix whose (i,j)th element is a_{ij}/b_ib_j. A total of 15 Fourier coefficients were chosen, the corresponding values of a_{ij} and b_i were calculated, and the 15×15 matrix was diagonalized with the aid of a computer.[4] The ground-state energy of the positron was found to be $+7.317$ eV.

V. POSITRONIUM FORMATION IN THE CRYSTAL

We consider now the energetics of positronium formation following the usual type of argument. Let E_p be the ground-state energy of the bound positronium system in the crystal, E^+ the actual energy of the positron, and $E_0{}^+$ the energy of its ground state, $-V_i$ the energy of an electron at the top of the valence band, and $-V_1$ the energy of an electron at the bottom of the conduction band.

In order for it to be energetically possible for a positron to capture an electron from the valence band to form positronium, the positron energy must be large enough to satisfy the inequality,

$$E^+-V_i\geqslant E_p. \quad (54)$$

In order for the positronium so formed to persist, energy considerations should forbid the separation of the bound system with the deposition of the electron at the bottom of the conduction band. Thus we require

$$E_0{}^+-V_1>E^+-V_i. \quad (55)$$

[4] The computation of the matrix elements and the diagonalization of the matrix was performed by A. G. Heinicke at the University of Manitoba.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Values of V_i and V_1 for LiH are not available, but for purposes of the present argument it will turn out that rough estimates of these quantities are sufficient. For the alkali halides, V_i ranges up to about 9 eV, and V_1 has been estimated of the order of 0.5 eV[5] or less. For use in (54) and (55) we shall set $V_i=10$ eV and $V_1=1$ eV. With the value of E_p obtained in Sec. III, (54) becomes

$$E^+ \geqslant -13.4+10=-3.4 \text{ eV}. \quad (56)$$

This is certainly satisfied even if we set E^+ equal to the value of $E_0^+=7.3$ eV obtained in Sec. IV.

It can, of course, be argued that the value used for E_0^+ is probably much too large, since in obtaining this value the positron was considered as moving in a fixed potential, no regard being taken of polarization effects. In the actual crystal the positron would be expected to concentrate the electron cloud about itself, with a resultant lowering of its energy. The amount by which the energy would thus be lowered would, however, not exceed the free positronium binding energy of 6.8 eV. Making allowance for this, even if we set $E^+=E_0^+=0$ in (55), the inequality is satisfied. Thus, the energy requirement for positronium formation is satisfied for positrons of energies down to the ground-state energy, E_0^+.

The inequality (55) can be rewritten

$$E^+-E_0^+<V_i-V_1\sim 9 \text{ eV}, \quad (57)$$

the right-hand side being now the measure of the width of the Ore gap. Thus, we have that a positron in the crystal with an energy anywhere within a range of some 9 eV above the ground-state energy can, by capture of an electron from the valence band, form a positronium-like bound system, which will be energetically stable against dissociation by deposition of the electron in the conduction band.

VI. DISCUSSION AND CONCLUSIONS

A number of questions arising out of the preceding treatment would seem to merit discussion at this point. These comprise the use of perturbation theory, the perhaps surprisingly low value for the upper bound to the positronium energy, the mean life for electron-positron annihilation from the bound state, and the neglect of exchange in treating the electron and positron of the bound system as being subject to the same effective crystal field.

Perturbation theory is used in the preceding treatment to the extent that it provides a wave function with which to calculate an upper bound to the positronium energy. It is not suggested that this function gives a particularly close approximation to the state of the system. The use of a better function would, however, result in a lower value of the upper bound, and this would strengthen the energetics argument of Sec. V. For the sake of completeness only, the above wave function is used in Appendix A to estimate the mean life of the positronium system, but obviously no great weight can be attached to the numerical result obtained.

The result obtained in Sec. III for the upper bound to the positronium energy might at first appear to be surprisingly low. It would be expected that the positronium in the crystal would suffer distortion such as to increase its internal energy; and this is indeed the case. However, positronium being an extended system, its constituent electron and positron will not have the same distribution in the crystal, with the result that the two particles may on the average occupy (different) positions of low respective potential energy. It is in fact the very distortion tending to raise the internal energy of the system which enables the positronium to "fit" the shape of the crystal field so as to lower its potential energy. An examination[6] of the expression (49) for the positronium energy reveals a very slow convergence of the series. This indicates the importance of the terms involving large quantum numbers. These higher terms do not greatly affect the internal energy, but through providing high angular momentum components in the wave function they give a distortion of the positronium which results in a large downward shift of the potential energy.

The use of the same form of crystal potential for both particles of the positronium system implies neglect of effects of exchange between the electron of the bound system and electrons in Bloch states of the crystal. An examination of this question, details of which are given in Appendix B, reveals that the exchange energy in the interaction between the electron of the positronium and an electron of the crystal is zero. This constitutes the justification for the neglect of exchange and for the use of the same form of crystal potential for the electron and the positron of the bound system.

The present investigation indicates that positronium formation is probably an important part of the process of positron annihilation in lithium hydride. Pending further investigation, one is tempted to speculate that this might also be the case in other ionic crystals such as other alkali hydrides and alkali halides. If positronium is formed in such a crystal in the triplet state, the subsequent two-photon annihilation of the positron with an electron through a "pick-off" process would be expected to contribute to a τ_2 component of the positron annihilation radiation. Such a τ_2 component has been observed[7] in a number of ionic crystals.

APPENDIX A: ANNIHILATION MEAN LIFE

For a calculation of the mean annihilation rate from the singlet state of the positronium-like system with

[5] A. J. Dekker, *Solid State Physics* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey), pp. 371–373.

[6] Details of the calculation are given in R. I. Verrall, M.Sc. thesis, University of Manitoba (unpublished).

[7] A. Bisi, A. Fiorentini, and L. Zappa, Phys. Rev. 131, 1023 (1963).

CC-0. In Public Domain. Gurukul

Digitized by Arya Samaj Foundation Chennai and eGangotri

state function ψ of Eq. (46), it is sufficient to calculate for this state the density of the electron at the location of the positron (averaged over all positions of the center of mass) and compare this with the corresponding density in the ground state of free positronium. The ratio of the two densities will then be equal to the ratio of the corresponding annihilation rates, that of free positronium being known.

It will be recalled that the terms involving excited states of positronium which occur in the expansion (46) all have $l>0$, and these terms vanish for $\mathbf{r}=0$. Thus, it is only the ground-state term, $\psi_{0,100}$, which contributes to the electron density desired, and the contribution of this term differs from that in free positronium only through the fact that ψ is not normalized to unity. Thus the electron density at the positron for positronium in the crystal is to the corresponding density in the ground state of free positronium in the ratio of $1/(\psi,\psi)=1/1.16$. Thus, for the wave function ψ the mean life of singlet positronium in the crystal is 1.16 times the mean life of singlet, ground-state, free positronium.

APPENDIX B: EXCHANGE EFFECTS

An examination will now be made of the nature of possible exchange effects, hitherto neglected. For this purpose we consider the system of positronium in the crystal in interaction with an electron in a Bloch state.

Let the wave function for positronium in the crystal be of the form of a superposition of free positronium functions similar to (46), i.e.,

$$\psi(1,2)=\sum A(\mathbf{k},nlm) \times\exp[\pi i\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)]\varphi_{nlm}(\mathbf{r}_1-\mathbf{r}_2), \quad (58)$$

where the subscripts 1 and 2 refer to the positron and electron, respectively. For the electron in the Bloch state we write the wave function (3). For the two systems in interaction we take a symmetrized function

$$\Psi=(1/\sqrt{2})[\psi(1,2)\chi(3)\pm\psi(1,3)\chi(2)]. \quad (59)$$

To examine the effect of exchange, we consider for definiteness the Coulomb interaction between the two electrons. Its average will be proportional to $(\Psi,(1/r_{23})\Psi)$. The direct term in this inner product is equal to

$$\int|\psi(1,2)|^2|\chi(3)|^2\frac{1}{r_{23}}d\tau, \quad (60)$$

while the exchange term is

$$\int\psi^*(1,2)\chi^*(3)\psi(1,3)\chi(2)(1/r_{23})d\tau. \quad (61)$$

In order to compare these two expressions, we select a typical term from the summation which results in each from the substitution of the expansion (58).

Out of the expansion of (60) we take the term

$$I=\int\exp[\pi i(\mathbf{K}'-\mathbf{K})\cdot(\mathbf{r}_1+\mathbf{r}_2)]\varphi_{nlm}^*(\mathbf{r}_1-\mathbf{r}_2) \times\varphi_{n'l'm'}(\mathbf{r}_1-\mathbf{r}_2)|\chi(\mathbf{r}_3)|^2d\tau. \quad (62)$$

By a coordinate transformation this can be rewritten

$$I=\int\exp[\pi i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{r}_1]\varphi_{nlm}^*(\mathbf{r}_1)\varphi_{n'l'm'}(\mathbf{r}_1)d\tau_1 \times\int\exp[2\pi i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{r}_2]\frac{1}{r_2}d\tau_2 \times\int|\chi(\mathbf{r}_3)|^2\exp[2\pi i(\mathbf{K}'-\mathbf{K})\cdot\mathbf{r}_3]d\tau_3. \quad (63)$$

Out of (61) we take the corresponding term

$$J=\int\exp[\pi i\{\mathbf{K}'\cdot(\mathbf{r}_1+\mathbf{r}_3)-\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)\}]\varphi_{nlm}^*(\mathbf{r}_1-\mathbf{r}_2) \times\varphi_{n'l'm'}(\mathbf{r}_1-\mathbf{r}_3)\chi(\mathbf{r}_3)\chi(\mathbf{r}_2)d\tau. \quad (64)$$

We can, by omitting the exponentials in (64), write

$$|J|\leqslant\int\varphi_{nlm}^*(\mathbf{r}_1-\mathbf{r}_2)\varphi_{n'l'm'}(\mathbf{r}_1-\mathbf{r}_3) \times\chi^*(\mathbf{r}_3)\chi(\mathbf{r}_2)\frac{1}{r_{23}}d\tau. \quad (65)$$

Because of the nature of the positronium functions φ the integral (65) converges, in spite of χ not being quadratically integrable. By contrast, the third integral in (63) diverges. Thus, a suitable box normalization to ensure the convergence of the direct interaction term I causes the exchange term J to vanish.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Oscillatory Magnetoresistance in Gray Tin*

E. D. Hinkley† and A. W. Ewald
Department of Physics, Northwestern University, Evanston, Illinois
(Received 10 January 1964)

The oscillatory component of magnetoresistance has been measured for single crystals of n-type gray tin at liquid-helium temperatures. The period of oscillation is independent of magnetic field direction, confirming an isotropic effective mass and, therefore, a conduction-band edge at the center of the Brillouin zone. Oscillations were also observed in a polycrystalline specimen. Carrier concentrations deduced from the oscillatory period are in reasonably good agreement with values obtained from the Hall coefficient over a range of almost two orders of magnitude in this parameter. An electron effective mass of 0.024 m_0 was obtained from the field and temperature dependence of the oscillatory amplitude for samples having carrier concentrations on the order of 10^{16} cm^{-3}. Nonthermal damping of the oscillations is attributed mainly to inhomogeneities in the impurity distribution. The experimental results are compared with theory.

I. INTRODUCTION

In recent years oscillatory magnetoresistance measurements have provided band-structure information for several of the compound semiconductors.[1–3] The results are in agreement with those obtained from cyclotron resonance[4–6] and de Haas–van Alphen[7] experiments. Of the *elemental* semiconductors, only germanium[8] and gray tin[9] have shown an oscillatory component of the magnetoresistance. The band structure of germanium had already been established through cyclotron resonance and other measurements before the magnetoresistance oscillations were observed, and therefore, this phenomenon was not exploited. In the case of gray tin, where cyclotron resonance has not been observed, low-temperature band-structure information was lacking. Previous magnetoresistance measurements at liquid-nitrogen temperature and above had indicated a change in symmetry type from that characteristic of [111] valleys at 273°K to the spherical symmetry of a central conduction-band minimum at 77°K (see Ref. 10). The present work indicates that at lower temperatures the electrons remain in the central minimum. It also yields a value for the effective mass of electrons in this band.

This paper describes the oscillatory magnetoresistance measurements made on single crystals (and on one polycrystalline specimen) of n-type gray tin. The experimental data are compared with magnetoresistance theories in terms of oscillatory period, phase, and amplitude. The value of electron effective mass and the shape of the constant energy surfaces for gray tin as deduced from the data are also compared with the theoretical predictions.

II. THEORY

When a magnetic field is applied to a system of charge carriers, the quasicontinuous distribution of states in momentum space collapses into a set of highly degenerate one-dimensional Landau subbands. Consequently, under certain conditions, the electrical resistance of a degenerate ($E_F > kT$) semiconductor is found to have an oscillatory dependence on magnetic field, periodic in reciprocal field. The oscillatory period is inversely proportional to the extremal cross-sectional area of the Fermi surface in a plane perpendicular to the magnetic field direction, and is generally a function of field orientation. For the special case of a spherical Fermi surface, however, the oscillatory period is a function only of carrier concentration.

$$P = (2e/\hbar c)(3\pi^2 n)^{-2/3} = 3.18 \times 10^6 n^{-2/3}, \qquad (1)$$

where n is measured in cm^{-3} and P in reciprocal gauss.

The usual semiquantitative conditions for the occurrence of quantum oscillations require that the separation $\hbar\omega$ between Landau levels be greater than the energy spread associated with the Fermi surface electrons. There are two major causes of this energy spread: First, a thermal "fuzzing" occurs at any non-

* This is a publication of the Materials Research Center, Northwestern University. It is based on a dissertation submitted by E. D. Hinkley to the Graduate School of Northwestern University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The work was supported jointly by the U. S. Office of Naval Research and by the Advanced Research Projects Agency of the U. S. Department of Defense.

† Present address: Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts.

[1] H. P. R. Frederikse and W. R. Hosler, Phys. Rev. **108**, 1136 (1957).

[2] R. J. Sladek, Phys. Rev. **110**, 817 (1958). H. P. R. Frederikse and W. R. Hosler, *ibid.* **110**, 880 (1958).

[3] K. F. Cuff, M. R. Ellett, and C. D. Kuglin, J. Appl. Phys. Suppl. **32**, 2179 (1961).

[4] G. Dresselhaus, A. F. Kip, C. Kittel, G. Wagoner, Phys. Rev. **98**, 556 (1955). E. Burstein, G. S. Picus, and H. A. Gebbie, *ibid.* **103**, 825 (1956).

[5] B. Lax and J. L. Mavroides, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1960), Vol. 11.

[6] U. J. Hansen and K. F. Cuff, Bull. Am. Phys. Soc. **8**, 246 (1963).

[7] P. J. Stiles, E. Burstein, and D. N. Langenberg, Bull. Am. Phys. Soc. **6**, 115 (1956).

[8] G. Lautz and W. Ruppel, Physik. Verhandl. **6**, 193 (1955). I. G. Fakidov and E. A. Zavadskii, Zh. Eksperim. i Teor. Fiz. **34**, 1036 (1958) [English transl.: Soviet Phys.—JETP **7**, 716 (1958)].

[9] E. D. Hinkley and A. W. Ewald, Bull. Am. Phys. Soc. **7**, 409 (1962). E. D. Hinkley and A. W. Ewald, Bull. Am. Phys. Soc. **8**, 245 (1963).

[10] O. N. Tufte and A. W. Ewald, Phys. Rev. **122**, 1431 (1961).

CC-0. In Public Domain. Gurukul

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Gray tin sample characteristics.

Sample	n_H (cm^{-3})[a]	$\sigma_{4.2}$ ($ohm^{-1}\,cm^{-1}$)	$\sigma_{4.2}/\sigma_{1.2}$	μ ($cm^2/V\text{-}sec$)[b]
A-1	4.17×10^{16}	772	1.00	1.16×10^{5}
B-1[c]	1.14×10^{16}	455	1.00	2.50×10^{5}
B-2[c]	...	465	1.00	...
B-3	...	375	1.00	...
B-5	1.28×10^{16}	...	...	...
B-6	0.82×10^{16}	...	...	...
B-7	1.76×10^{16}	...	...	...
C-1	4.3×10^{15}	240	1.04	3.49×10^{5}
C-2	6.6×10^{15}	240	1.02	2.28×10^{5}
D-2	6.4×10^{14}	63	4.5	6.2×10^{5}

[a] Obtained from Hall coefficient, R, at 4.2°K as $1/Re$.
[b] Calculated as the product $R\sigma$ at 4.2°K.
[c] Samples obtained from same single crystal.

zero temperature since the electrons may deviate from the Fermi surface by a few kT; second, collisions limit the lifetime τ of a quantum state, with concomitant uncertainty $\hbar/\tau$ in electron energy leading to the cyclotron condition $\omega\tau>1$. This condition may be written more conveniently as $\mu B>10^8$, where μ is the Hall mobility in cm^2/V sec and B is the magnetic field in gauss.

In addition to the thermal and collision broadening of the electron energy levels described above, the Fermi level itself may vary throughout the specimen as a result of an inhomogeneous distribution of impurities. It is clear from Eq. (1) that such an effect would be manifested as a spatial variation of the oscillatory period, and partial cancellation of oscillations would occur.

The above conditions, together with their implications concerning the observation of quantum oscillations at readily obtained steady magnetic fields, are summarized below:

(1) $\hbar\omega>kT$. Low carrier effective mass, low ambient temperature.
(2) $\mu B>10^8$. High carrier mobility.
(3) $\hbar\omega>\Delta E_F$. Homogeneous specimen.

A recent theory for transverse ($B\perp I$) oscillatory magnetoresistance proposed by Adams and Holstein[11] embodies all the above conditions, except for inhomogeneity broadening. The latter is taken into account by defining a combined nonthermal broadening temperature T' as the sum of the collision (Dingle) and inhomogeneity broadening temperatures. The analysis of Adams and Holstein, while limited to materials with spherical Fermi surfaces, shows the oscillatory phenomenon to be essentially independent of the type of scattering. On the basis of their derivation, the oscillatory part of the transverse magnetoresistance of interest here may be written as

$$\frac{\Delta\rho_{0sc}}{\rho_0}=5\sqrt{2}\pi^2\frac{kT}{E_F^{1/2}(\hbar\omega)^{1/2}}\sum_{M=1}^{\infty}\frac{M^{1/2}(-1)^M e^{-2\pi^2M(kT'/\hbar\omega)}}{\sinh[2\pi^2M(kT/\hbar\omega)]}\times\cos\left[\frac{2\pi ME_F}{\hbar\omega}-\frac{\pi}{4}\right], \quad (2)$$

where ρ_0 is the zero-field resistivity and M is the harmonic of the oscillations. This expression may be simplified because only the fundamental ($M=1$) is usually significant. In addition, the hyperbolic sine may be replaced by its exponential approximation. Equation (2) then reduces to

$$\frac{\Delta\rho_{0sc}}{\rho_0}=-10\sqrt{2}\pi^2\frac{kT}{E_F^{1/2}(\hbar\omega)^{1/2}}\times\exp\left[-\left(\frac{2\pi^2k}{\beta_0}\right)\left(\frac{m^*}{m_0}\right)\left(\frac{T+T'}{B}\right)\right]\cos\left(\frac{2\pi}{PB}-\frac{\pi}{4}\right), \quad (3)$$

where the oscillatory period P, the effective mass m^*, and $\beta_0(=e\hbar/m_0c$, the double Bohr magneton) have been introduced.

Oscillations in the longitudinal ($B||I$) magnetoresistance have been treated by Argyres.[12] The corresponding equation is identical with Eq. (3) when the latter is reduced by a factor of 5.

III. EXPERIMENTAL DETAILS

A. Specimens

Gray tin crystals were grown from mercury solution by the procedure of Ewald and Tufte.[13] Crystals of various impurity concentrations were available as a result of different purification processes carried out on the commercial zone-refined tin. The different purities were reflected in different carrier concentrations at 4.2°K as determined by the Hall coefficient. The samples ranged in carrier concentration from 6.4×10^{14} cm^{-3} to 4.2×10^{16} cm^{-3} and on this basis were divided into four categories designated by a letter prefix to the sample number as shown in Table I. The donor impurity atoms are probably antimony which is not removed during the zone-refining process, and only partially removed during subsequent purification. Cutting of the crystals into rectangular parallelepipeds suitable for electrical measurements was performed with a sand-blast unit. Sample dimensions were approximately $0.2\times0.4\times3$ mm.

Samples were mounted on Bakelite holders at the end of long, thin-walled Monel tubes. Current electrodes, which also served as support for the sample, were of No. 36 copper wire and were attached to the ends of the

[11] E. N. Adams and T. D. Holstein, Phys. Chem. Solids 10, 254 (1959).
[12] P. N. Argyres, Phys. Chem. Solids 4, 19 (1958).
[13] A. W. Ewald and O. N. Tufte, J. Appl. Phys. 29, 1007 (1958)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

sample with gallium-indium (3 parts Ga to 1 part In) solder. To reduce adverse effects due to large area potential-probe contacts, the following soldering procedure was employed. Bare No. 56 (0.0127 mm) copper wires, which served as the potential probes, were first "tinned" and then brought into contact with the cooled sample at the desired locations. A 200 mA current pulse was then passed through the potential leads producing localized heating which melted the solder. By this means the solder contact area was held to less than that of the wires. The potential probes were usually attached to the side of the sample and (to eliminate "edge" effects) at a distance from the ends greater than the sample width. In some instances, however, the potential probes were mounted at the ends of the specimens with the current electrodes.

B. Apparatus and Procedure

The sample was immersed directly in liquid helium. Temperatures below 4.2°K were obtained by pumping on the helium vapor, and were measured by the vapor pressure. The lower tip of the helium Dewar was located between the pole pieces of a magnet which could be rotated in the horizontal plane. Magnet current was supplied by a bank of storage batteries, and regulated by water-cooled rheostats.

A constant current of up to 10 mA was supplied to the sample by a high-impedance transistorized current source. Maximum sample current was determined by the condition that the rise in sample temperature as a result of joule heating should not be excessive. Analysis of the oscillatory amplitudes for a sample under varying conditions of power dissipation showed that if the input power was less than 100 μW, sample temperature rose by less than 0.1°K.

The amplified voltage from the potential probes was applied to the Y axis of an X-Y recorder, the X axis of which was driven by a signal proportional to the current through the magnet. Magnetic field values corresponding to various positions along the X axis were obtained from a calibration curve made with a rotating-coil fluxmeter. In those cases in which the nonoscillatory part of the magnetoresistance was approximately linear with field, a bucking voltage proportional to magnet current was used to permit greater amplification of the oscillations. For precise measurements, e.g., to determine Hall voltage or zero-field resistance, the electrometer amplifier was replaced by a potentiometer.

IV. DATA REDUCTION TECHNIQUE

A. Oscillatory Period and Phase

The period and phase of the magnetoresistance oscillations were usually obtained from a direct analysis of the X-Y recorder curves. (In order to determine the nonoscillatory magnetoresistance, however, the data had to be averaged over positive and negative field directions to cancel Hall voltage contributions resulting from misalignment of the potential probes.) Envelope curves were constructed tangent to the oscillations with a nodal line drawn halfway between. The intersections of the nodal line with the recorder curve located the nodes of the oscillations which were used to determine the period and phase.

Equation (3) may be expressed in terms of an amplitude A, period P, and phase ϕ as

$$\Delta\rho_{0sc}/\rho_0 = A(B,T)\cos[(2\pi/PB+\phi+\pi], \quad (4)$$

where the minus sign has been absorbed into the argument of the cosine. A node occurs whenever the argument of the cosine is an odd integer multiple of $\pi/2$. This condition may be stated as

$$(2\pi/PB_n)+\phi+\pi=(2n+\tfrac{1}{2})\pi, \quad (5)$$

where n takes on both integer and half-integer values corresponding to the two nodes per oscillation. Ambiguity in the assignment of integers and half-integers to the experimentally observed nodes is removed by noting that, according to Eqs. (4) and (5), the oscillatory curve (as a function of B) has positive slope at the integer nodes and negative slope at the half-integer nodes. This point is very important to an experimental determination of phase. A plot of reciprocal magnetic field values at which the nodes occur versus integers and half-integers represents a straight line of slope P. The intercept n_0 with the $1/B=0$ axis is related to the phase. These relationships become clear upon rearranging Eq. (5) into the following form:

$$1/B_n = Pn-[(\phi/2\pi)+\tfrac{1}{4}]P. \quad (6)$$

The phase can be deduced from n_0 by the expression

$$\phi=(2n_0-\tfrac{1}{2})\pi. \quad (7)$$

If the measured phase agrees with the predicted value of $-\frac{1}{4}\pi$, then $n_0=\frac{1}{8}$. For future reference, an intercept $n_0=0$ implies a phase of $-\frac{1}{2}\pi$. Since integral multiples of 2π may be added algebraically to ϕ without altering the situation, n_0 is likewise arbitrarily defined with respect to the integers.

B. Effective Mass

Equation (1) shows that, for a semiconductor having spherical constant-energy surfaces, the oscillatory period is a function only of carrier concentration, and is, therefore, independent of the effective mass. The effective mass parameter is, however, contained in the oscillatory amplitude expression,

$$A(B,T)\propto B^{-1/2}\exp[-\gamma(m^*/m_0)(T+T')/B], \quad (8)$$

where $\gamma\equiv 2\pi^2k/\beta_0=14.68\times10^4$ gauss/deg If the above equation is multiplied by $B^{1/2}$, and the logarithm of the resulting expression differentiated with respect to $1/B$,

CC-0. In Public Domain. Gurukul

Digitized by Arya Samaj Foundation Chennai and eGangotri

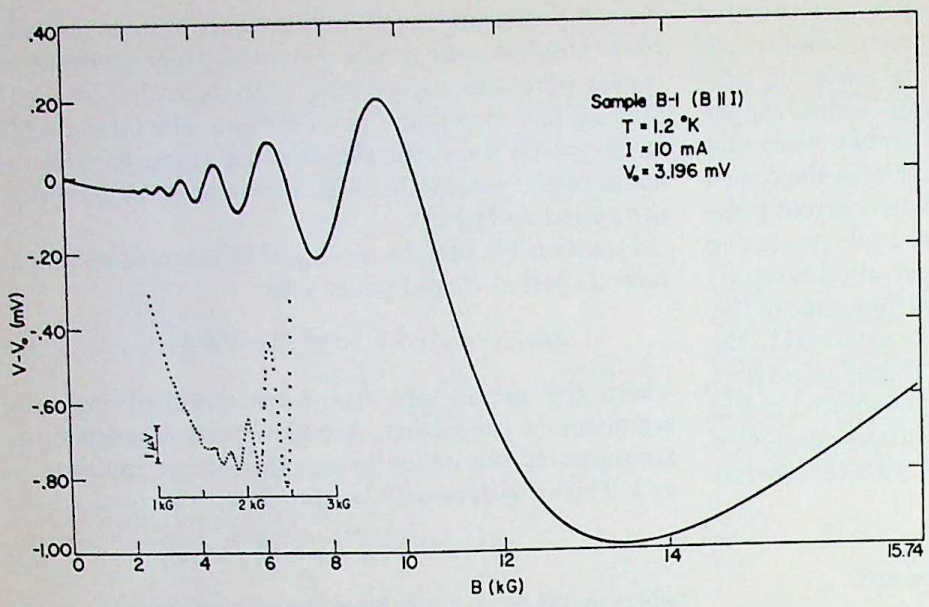


FIG. 1. Recorder traces of magnetoresistance voltage versus magnet current for gray tin single-crystal sample B-1 at 1.2°K. Data in the low-field region (inset) were recorded point-by-point because the signal was comparable to induced voltages occurring when the magnetic field was changed. Magnetic field is parallel to sample current.

the following is obtained:

$$d\ln(AB^{1/2})/d(1/B) = -\gamma(m^*/m_0)(T+T'). \quad (9)$$

Therefore, a logarithmic plot of $AB^{1/2}$ versus $1/B$ yields a straight line, the slope of which is proportional to the product $m^*(T+T')$. The unknown quantities m^* and T' may be evaluated by analyzing the field dependence of the oscillatory amplitude for at least two different ambient temperatures. For some of the samples tested *several* ambient temperatures were used in an attempt to overdetermine the effective mass value and (hopefully) increase its accuracy. We found it convenient to define a slope function $S(T)$ equal to the negative of Eq. (9). The points of an S versus T plot represent a straight line having slope $\gamma(m^*/m_0)$ and an intercept (at the $T=0$ axis) of $\gamma(m^*/m_0)T'$.

V. RESULTS AND DISCUSSION

A. Oscillatory Parameters

Figure 1 shows actual recorder traces of voltage across the potential probes of sample B-1 as a function of magnet current. The continuous curve exhibits oscillations starting somewhat above 2 kG. Oscillations at still lower fields are apparent in the inset curve which was taken with one hundred times greater amplification. The latter curve was recorded point-by-point to eliminate traces of induced voltages due to the quasi-continuous change in magnetic field.

Since sample B-1 exhibited the largest number of detectable oscillations of all those tested, it was of special interest to our investigation. Figures 2 and 3 show the transverse and longitudinal magnetoresistance field dependences, respectively, for this sample. It is noteworthy that the oscillatory amplitudes for the two modes are approximately the same, although an order of magnitude smaller than predicted by Eq. (3). The nonoscillatory magnetoresistance components differ

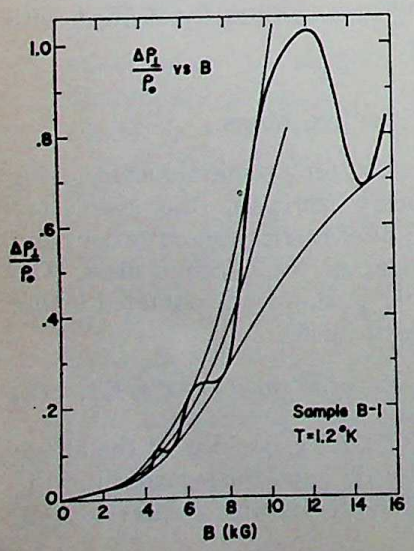


FIG. 2. Field dependence of transverse magnetoresistance for sample B-1 at 1.2°K. This curve was obtained by averaging data for positive and negative field directions. A line of nodes has been constructed midway between envelope lines, and was used to determine the oscillatory period and phase.

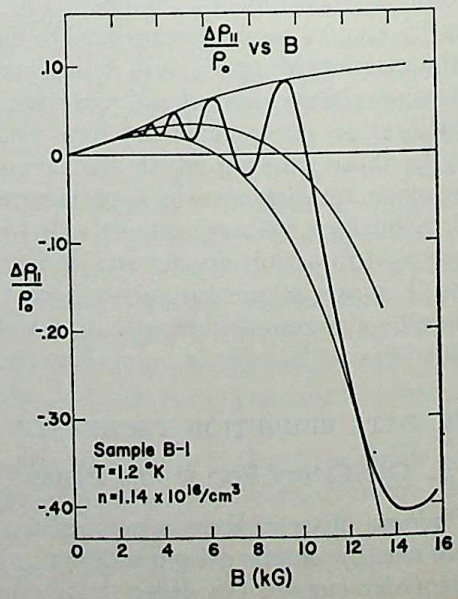


FIG. 3. Field dependence of longitudinal magnetoresistance for sample B-1 at 1.2°K. Negative longitudinal magnetoresistance was observed for all samples in the high-field region.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

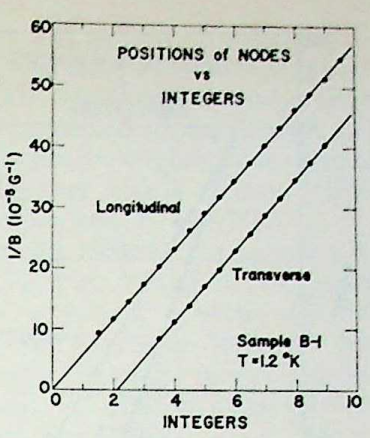


FIG. 4. Nodal positions of transverse and longitudinal magnetoresistance oscillations for sample B-1 are plotted versus integers and half-integers. The oscillatory periods and phases were obtained from the slopes and intercepts, respectively, of the straight lines.

markedly between the transverse and longitudinal modes, and the *negative* longitudinal magnetoresistance apparent at high fields was observed at liquid-helium temperatures for all samples. Node-integer plots for both orientations of magnetic field are shown in Fig. 4. The oscillatory periods deduced from the slopes in accordance with Eq. (6) are 5.84×10^{-5} G^{-1} for the transverse mode and 5.72×10^{-5} G^{-1} for the longitudinal mode. Since this sample was mounted so that the magnetic field for both modes was [110] oriented, this agreement is expected.

A detailed study of the dependence of oscillatory period on crystal orientation was made using sample B-6. Figure 5 shows the oscillatory period for this sample as a function of magnetic field direction. If we attribute the 3% scatter of points to errors in the data reduction, we conclude that the period is independent of crystal orientation. This *indifference* of oscillatory period to magnetic field direction relative to the crystal axes was found on all samples and is the basis for concluding that, at liquid helium temperatures, the electrons occupy a spherically symmetric band.

The absence of a crystal orientation dependence of the oscillatory period was also verified by the method

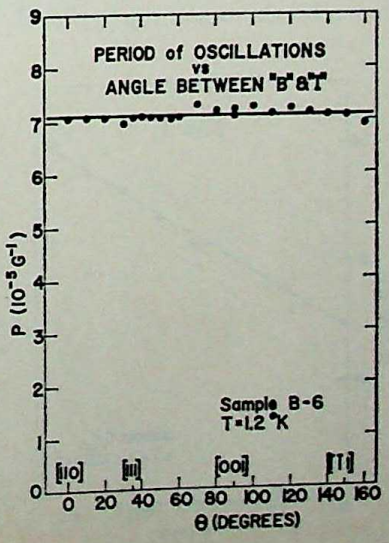


FIG. 5. Angular dependence of oscillatory period for sample B-6. Sample current is in the [110] direction, with magnetic field in the (1$\bar{1}$0) plane. The 3% scatter of points from the average value is attributed to errors in the graphical data reduction procedure.

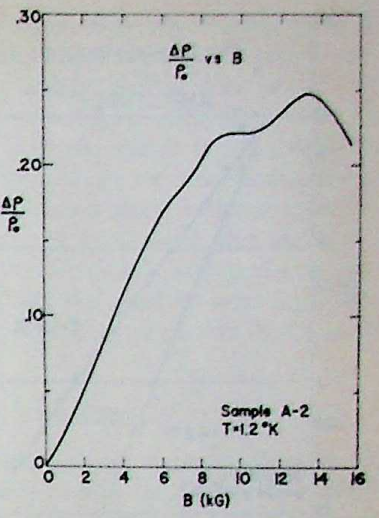


FIG. 6. Field dependence of magnetoresistance for polycrystalline sample A-2. The presence of oscillations confirms the orientation-independence of oscillatory period.

of Sladek[2] which utilizes a polycrystalline specimen. A sample was prepared by transforming a thin white-tin foil to the gray phase. (X-ray analysis has shown that samples prepared in this manner are highly polycrystalline.) The magnetoresistance field dependence for this sample is shown in Fig. 6, where the presence of oscillations confirms sphericity of the constant energy surfaces.

TABLE II. Results of oscillatory analysis.

Sample	n_H (cm^{-3})	n_P (cm^{-3})	m^*/m_0	T' (°K)[a]
A-1	4.17×10^{16}	2.16×10^{16}	...	...
B-1[b]	1.14×10^{16}	1.30×10^{16}	0.024[c]	3.2
B-2[b]	...	1.45×10^{16}	0.025	5.4
B-3	...	1.07×10^{16}	0.021	9.3
B-5	1.28×10^{16}	0.97×10^{16}	0.023[c]	5.3
B-6	0.82×10^{16}	0.96×10^{16}	0.022	5.4
B-7	1.76×10^{16}	1.28×10^{16}	...	...
C-1	4.3×10^{15}	4.2×10^{15}	0.025[c]	3.9
C-2	6.6×10^{15}	5.0×10^{15}	...	...
D-2	6.4×10^{14}	5.7×10^{14}	...	...

[a] Combined nonthermal broadening temperature.
[b] Samples obtained from same single crystal.
[c] Effective mass value determined on the basis of seven different ambient temperatures. Others on basis of two different ambient temperatures.

Equation (1) may now be used to compare the carrier concentration n_P as measured by the oscillatory period,

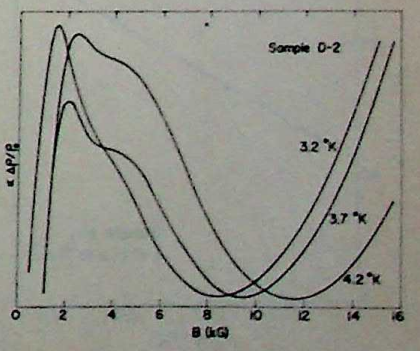


FIG. 7. Field dependence of magnetoresistance for pure sample D-2. The increase of oscillatory period with a decrease in temperature is caused by carrier freeze-out.

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

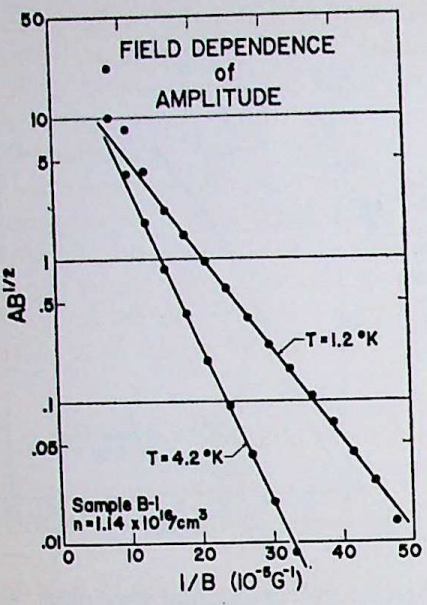


FIG. 8. Field dependence of oscillatory amplitude for sample B-1 at ambient temperatures of 4.2 and 1.2°K. Slopes are proportional to the product $m^*(T+T')$.

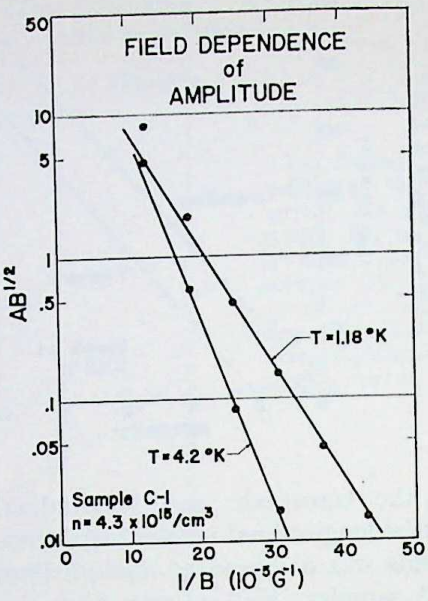


FIG. 10. Field dependence of oscillatory amplitude for sample C-1 at ambient temperatures of 4.2 and 1.18°K. Smaller number of data points (cf. Fig. 8) reflects longer period of this higher purity specimen.

with that given by the Hall coefficient. For sample B-1, $n_P=1.30\times10^{16}$ cm^{-3}, whereas $n_H=1.14\times10^{16}$ cm^{-3}. Similar comparisons for other samples are shown in Table II. It should be noted that the n_H value for any particular sample is probably somewhat dependent on the location of the Hall probes because of inhomogeneity in the impurity distribution.

Small magnetoresistance oscillations were detected in the very pure sample D-2 at 4.2°K. The large decrease in conductivity for this sample from 4.2 to 1.2°K (cf. Table I) led to a search for a temperature dependence of the oscillatory period, indicative of carrier freeze-out. Figure 7 illustrates the magnetoresistance field dependence for sample D-2 at three different ambient temperatures. The period is very long (corresponding to a carrier concentration of only 6.4×10^{14} cm^{-3}), but a change of period with temperature is quite apparent. A shifting of the extrema toward smaller magnetic field values with decreasing temperature indicates an increasing period, in qualitative agreement with Eq. (1).

For a discussion of oscillatory phase, we return to Fig. 4. The intercepts n_0 for the transverse and longitudinal modes are $\frac{1}{8}$ and 0, respectively, which correspond to phases of $-\frac{1}{4}\pi$ and $-\frac{1}{2}\pi$. Thus, the measured *transverse* phase agrees with theory, but a significant discrepancy exists for the *longitudinal* mode (predicted to be $-\frac{1}{4}\pi$ also). Several further measurements were made on this particular sample over a period of a few months (between which the sample was removed from its holder), and the above phases were consistently obtained. Moreover, a transverse phase of $-\frac{1}{4}\pi$ was observed on all other samples as well.

B. Electron Effective Mass Determination

The magnetic field dependence of the oscillatory amplitudes for sample B-1 at ambient temperatures of 1.2 and 4.2°K is shown in Fig. 8. [Deviations from linearity at the high-field region stem, at least in part, from a breakdown of the hyperbolic sine approximation

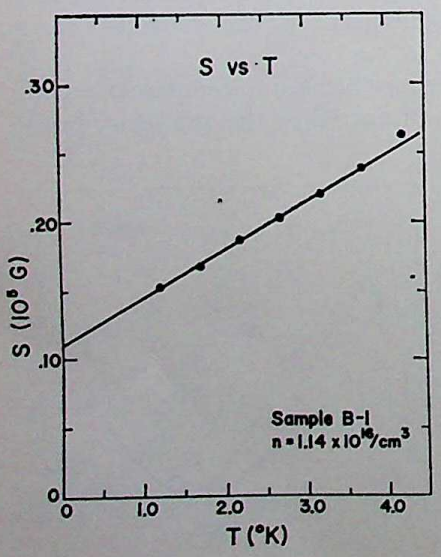


FIG. 9. Temperature dependence of slope function for sample B-1. Slope of line through points equals $\gamma(m^*/m_0)$. Intercept with ordinate equals $\gamma(m^*/m_0)T'$.

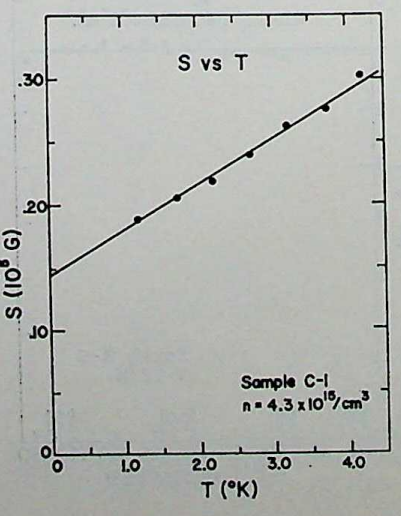


FIG. 11. Temperature dependence of slope function for sample C-1.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of Eq. (3).] Five other intermediate temperatures at half-degree intervals were also used, but these data were omitted for clarity. In Fig. 9 are shown the slopes $S(T)$ for all seven ambient temperatures. From the slope and intercept of this straight line an effective mass value of $0.024m_0$ was found, together with a combined nonthermal broadening temperature, T', of 3.2°K.

A similar procedure was followed for sample C-1, with the results shown in Figs. 10 and 11. Because of the fewer detectable oscillations exhibited by this higher purity sample, the results are somewhat less accurate, as is evidenced by the larger spread of points in Fig. 11. Nevertheless, an effective mass value of $0.025m_0$ was found, with $T'=3.9$°K.

In Table II are listed the experimentally obtained values for the electron effective mass in gray tin. Although T' varies from 3.2° to 9.3°K, the mass values themselves are fairly consistent. As mentioned previously, T' is the sum of the Dingle and inhomogeneity broadening temperatures. Because of the high mobility of these specimens, the Dingle temperature is less than 1°K (cf. Ref. 2). Consequently, we conclude that the major contribution to T' is from inhomogeneities in the specimens.

The spherical symmetry of the conduction band observed experimentally and discussed in the previous section was predicted by Herman,[14] and Liu and Bassani.[15] However, the calculated effective mass value of $0.0036m_0$ (see Ref. 16) which is based in part on the experimental band gap, is almost an order of magnitude smaller than our most accurate value of $0.024m_0$. Correction of the experimental value for nonparabolicity of the band[17] reduces it only slightly, to $0.020m_0$. To account for the relatively large observed mass as well as the pressure behavior of the conductivity at temperatures below 160°K, Groves and Paul[18] have recently proposed a new model for the band structure of gray tin. In this the direct gap is fixed at zero at all temperatures, and the temperature dependence of the conductivity in the intrinsic range is attributed to transitions from the uppermost valence band edge to [111] valleys. While a detailed quantitative comparison between the predictions of this model and all experimental data has not yet been made, the model appears to be more successful than the usual nonzero gap model in accounting for the experimental results in the lower temperature range.

VI. CONCLUSION

An oscillatory component of magnetoresistance has been observed at liquid-helium temperatures in single crystals of n-type gray tin. Exceptions to the general agreement between the experimental results and the theories of Adams and Holstein, and Argyres are the smaller observed amplitude and a measured longitudinal phase of $-\frac{1}{2}\pi$.

The *indifference* of oscillatory period to magnetic-field orientation relative to the crystal axes indicates that the minimum of the conduction band is at the center of the Brillouin zone, in agreement with the theoretical predictions of Herman, and Liu and Bassani. This was also confirmed by detecting oscillations in a polycrystalline specimen. A dependence of the oscillatory period on temperature observed in a very pure specimen was attributed to carrier freeze-out.

On the basis of the field and temperature dependence of the oscillatory amplitude we have obtained an electron effective mass value of $0.024m_0$, which does not appear to depend on carrier concentration. This is much larger than the value $0.0036m_0$ calculated by Bassani and Liu, and suggests that a major revision of the band structure model, such as that recently proposed by Groves and Paul, is required.

[14] F. Herman, J. Electron. 1, 103 (1955).
[15] L. Liu and F. Bassani, Bull. Am. Phys. Soc. 8, 51 (1963).
[16] F. Bassani and L. Liu, Phys. Rev. 132, 2047 (1963).
[17] J. Kolodziejczak, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Academic Press Inc., New York, 1961), p. 950.
[18] Steven Groves and William Paul, Phys. Rev. Letters 11, 194 (1963).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

$FeAl_2O_4$–$MgAl_2O_4$: Growth and Some Thermal, Optical, and Magnetic Properties of Mixed Single Crystals

GLEN A. SLACK
General Electric Research Laboratory, Schenectady, New York

(Received 30 December 1963)

Synthetic single crystals of $FeAl_2O_4$ have been grown from the melt. Measurements of the thermal conductivity K and optical absorption coefficient α have been made on a combination of synthetic and natural single crystals of $Mg_{1-x}Fe_xAl_2O_4$ for $0 \leq x \leq 1$. The α measurements at 300°K in the absorption band of tetrahedral Fe^{2+} between 0.4 and 0.6 eV have been calibrated and then used to measure x. The thermal conductivity results between 3 and 300°K show that the heat is conducted by phonons and that tetrahedral Fe^{2+} is a very effective phonon scatterer at all temperatures because it possesses a series of low-lying energy levels in the $3d$ shell. This type of scattering is absent in $CoAl_2O_4$. The phonon scattering by Fe^{2+} is dominant for $x > 0.1$, and for these crystals $K \sim T/\sqrt{x}$. The magnetic susceptibility χ of a single crystal of $FeAl_2O_4$, measured between 2 and 290°K, shows a peak at 9.5±0.5°K. The analysis of the χ and K measurements suggests that the ~10°K transition is caused by the individual Fe^{2+} ions, and is not associated with a long-range, antiferromagnetic ordering of the $FeAl_2O_4$.

INTRODUCTION

THE simplest magnetic analog of the basic diamagnetic spinel $MgAl_2O_4$ is one in which all or part of the Mg is replaced by a transition-metal ion M. The present series of studies concerns crystals in which M is divalent Fe. The compound $FeAl_2O_4$ is a normal spinel[1,2] as is[3,4] $MgAl_2O_4$ in which almost all of the Fe (or Mg) ions are in the tetrahedral (or A) lattice sites. The magnetic properties in a completely normal spinel are thus determined by the magnetic A-A interaction. Since the A-A separation is 3.52 Å in $FeAl_2O_4$, this interaction is very weak.[5] Thus, it is not possible for $FeAl_2O_4$ to become magnetically ordered until very low temperatures (~ 4°K) are reached. The following measurements at and below 300°K of the thermal conductivity, optical absorption, and magnetic susceptibility of synthetic single crystals of $FeAl_2O_4$ and of natural mixed crystals of $Mg_{1-x}Fe_xAl_2O_4$ give some information on the behavior of the individual Fe ions and some on the nature of the A-A interaction.

SINGLE CRYSTALS

The samples used in the present study were a combination of synthetic and natural crystals. See Table I. High-purity oligocrystalline $MgAl_2O_4$ prepared[6] by Navias by the vapor reaction of MgO with Al_2O_3 was used as a comparison standard. A large sample 1.2×0.3×0.3 cm composed of about five single crystals was obtained, and was used for the thermal conductivity and optical absorption measurements. A combination of chemical and x-ray analyses indicated this sample, R68, was a nearly stoichiometric spinel with a composition, 1 MgO+1.02 Al_2O_3, slightly rich in Al_2O_3. The crystallite diameter of 0.3 cm is sufficiently large so that these results should be representative of single-crystal values.

Natural crystals of $Mg_{1-x}Fe_xAl_2O_4$, where $0.0005 \leq x \leq 0.3$ were obtained through various sources. They originated in Australia (Queensland), Burma, and Ceylon. For the lowest values of x they were light pink in color. This color is caused by the presence of Cr in these, as in almost all natural spinel samples. For $x = 0.3$, the crystals were opaque black in the massive state, but in thin, polished sections were a yellowish-brown color.

No good natural crystals of hercynite, $FeAl_2O_4$, were available. Therefore synthetic single crystals of $FeAl_2O_4$ were grown from the melt in an iridium crucible under a controlled partial pressure of oxygen.

$FeAl_2O_4$ Crystals

Since single crystals of $FeAl_2O_4$ do not appear to have been grown before, it seems worthwhile to describe their preparation. The FeO-Al_2O_3 phase diagram at high temperatures has been worked out by Fischer and

TABLE I. A list of the single crystals of $Mg_{1-x}Fe_xAl_2O_4$ used in the present studies.

Sample number	a_0 Å	x[a]	Source	Color
R68	8.082	3×10^{-5}	synthetic	colorless in bulk
R42	8.0866	5×10^{-4}	Burma	light pink in bulk
R54	8.0868	6.0×10^{-4}	Burma	light pink in bulk
R97	8.089	2.4×10^{-3}	Ceylon	bluish purple in bulk
R56	8.128	1.9×10^{-1}	Australia	opaque, yellowish-brown thin sections
R62	8.117	2.7×10^{-1}	Australia	opaque, yellowish-brown thin sections
R67	8.138	~1.0	synthetic	opaque, greenish-yellow in thin sections
R75	8.140	~1.0	synthetic	opaque, greenish-yellow in thin sections

[a] x = fraction of the A sites occupied by Fe^{2+} ions as determined from the 0.6 eV optical absorption peak. For sample R68 the Fe concentration was determined from an emission spectrograph.

[1] T. W. F. Barth and E. Posnjak, Z. Krist. 82, 325 (1932).
[2] F. C. Romeijn, Philips Res. Rept. 8, 304, 321 (1953).
[3] G. E. Bacon, Acta Cryst. 5, 684 (1952).
[4] S. Hafner and F. Laves, Z. Krist. 115, 321 (1961).
[5] W. Roth, Proceedings of the International Colloquium of Neutron Diffusion and Diffraction, Grenoble, 1963 (unpublished); J. Phys. Radium (to be published).
[6] L. Navias, J. Am. Ceram. Soc. 44, 434 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Hoffmann[7] and by Galakhov.[8] They show that $FeAl_2O_4$ is the only ternary compound in the system, and it has a congruent melting point between 1780 and 1800°C. The partial pressure of oxygen at the phase boundary where $FeAl_2O_4$ dissociates into Fe, Al_2O_3, and O_2 has been measured[9–12] between 900 and 1100°C. When these data are extrapolated to the melting point at 1780°C, the predicted dissociation pressure is 10^{-7} atm of oxygen. A gas mixture of carbon monoxide to carbon dioxide in a molecular ratio of 8 to 1 at a total pressure in the vicinity of 1 atm will possess[13] this partial pressure of oxygen at 1780°C. Such an 8 to 1 mixture will provide an oxygen pressure about 5 times the dissociation pressure at 1000°C. Thus, a CO/CO_2 gas mixture of constant composition is a nearly ideal medium to use for growing $FeAl_2O_4$ since it will supply an oxygen partial pressure very close to the one needed. If the oxygen partial pressure becomes too high, much of the Fe^{2+} will be converted to the unwanted Fe^{3+}. The major problem in making $FeAl_2O_4$ is to keep as much of the iron as possible in the divalent Fe^{2+} state. In analogy with wüstite, "FeO," it may never be possible to have all of the iron in the divalent state.[14]

The $FeAl_2O_4$ crystals were therefore grown in a CO/CO_2 atmosphere with a molecular ratio of 1 part CO to 1 part CO_2. This is slightly more oxidizing than necessary in order to provide a margin of safety. Following the work of Horn[15] on ferrites, crucibles of pure iridium were used. These were in the shape of a tapered cylinder with a conical end to facilitate growth of only a few nuclei. The crystal growing furnace is shown in Fig. 1. The iridium crucible is placed inside of an iridium cylinder with a 0.05 cm wall thickness which acts as an rf susceptor. The heat is supplied by a 0.5-Mc, 25-kW radiofrequency oscillator. The rf coil, 20 cm long, is wound with a change in diameter and pitch at its midpoint, as shown in Fig. 1. The operating power level is adjusted so that the $FeAl_2O_4$ in the crucible is liquid when the crucible is in the upper part of the rf coil, and is solid when in the lower part. The temperature difference between the two parts of the coil is approximately 300°C. The motion of the temperature gradient in this modified Bridgman technique is produced by moving the rf coil upward past the crucible at a rate of 1 cm/h. After the melt has solidified, the power level of the rf oscillator is slowly reduced to zero over a period of hours. The thermal insulation around the iridium susceptor was composed of zirconia ceramic, Pt-40% Rh radiation shields, alumina ceramic, and quartz. Most of this detail has been omitted from Fig. 1. for clarity.

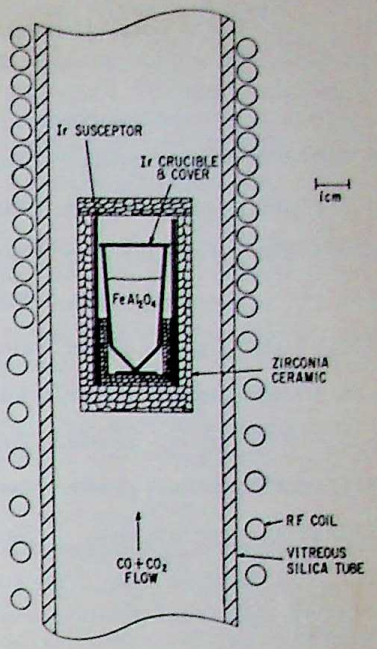


FIG. 1. A schematic of the apparatus used for growing single crystals of $FeAl_2O_4$ by a modified Bridgman technique. The rf coil is moved upward past the stationary crucible.

Large, sound single crystals of the order of 1 cm^3 in volume were produced by this technique. Metallographic and x-ray analyses indicated a single-phase material with the spinel-crystal structure and a lattice parameter a_0 of 8.138±0.001 Å for the as-grown sample R67. This value for a_0 is in reasonable agreement with the literature values[5,16–22] listed in Table II. It appears to be somewhat smaller than the most recent values on samples equilibrated at 1000°C, which cluster around 8.151 Å. If the results of Holgersson[16] on a natural crystal are corrected for the small concentration of $MgAl_2O_4$ in his sample, one obtains 8.155±0.004 Å. The variation in a_0 of a nominally pure $FeAl_2O_4$ sample can be caused[20,22] by the formation of mixed crystals with Al_2O_3, which reduces a_0, or with Fe_3O_4, which increases a_0. It can also be caused by a partial inversion[5,22] of the spinel from its ideal normal structure, or by oxidation of some Fe^{2+} to Fe^{3+}.

A chemical analysis of the as-grown crystal R67

[7] W. A. Fischer and A. Hoffmann, Arch. Eisenhüttenw. 27, 344 (1956).
[8] F. Ya. Galakhov, Izvest. Akad. Nauk SSSR, Otd. Khim. Nauk (1957), p. 525.
[9] R. Schenk, H. Franz, H. Willeke, Z. Anorg. Allgem. Chem. 184, 1 (1929).
[10] V. Cirilli, Gazz. Chim. Ital. 76, 339 (1946).
[11] F. D. Richardson, J. H. E. Jeffres, G. Withers, J. Iron Steel Inst. (London) 166, 213 (1950).
[12] R. G. Richards and J. White, Trans. Brit. Ceram. Soc. 53, 233 (1954).
[13] F. D. Richardson and J. H. E. Jeffres, J. Iron Steel Inst. (London) 160, 261 (1948).
[14] L. S. Darken and R. W. Gurry, J. Am. Chem. Soc. 67, 1398 (1945).
[15] F. H. Horn, J. Appl. Phys. 32, 900 (1961).
[16] S. Holgersson, Lunds Univ. Arsskr. Adv. 2 38, 1 (1927); Chem. Abstr. 24, 804 (1930).
[17] G. L. Clark, A. Ally, A. E. Badger, Am. J. Sci. 22, 539 (1931).
[18] O. Krause and W. Thiel, Z. Anorg. Allgem. Chem. 203, 120 (1931).
[19] H. Tazaki and S. Kuwabara, J. Sci. Hiroshima Univ. A15, 263 (1952).
[20] A. Hoffmann and W. A. Fischer, Z. Physik. Chem. (Frankfurt) 7, 80 (1956).
[21] L. M. Atlas and W. K. Sumida, J. Am. Ceram. Soc. 41, 150 (1958).
[22] A. C. Turnock, Carnegie Inst. Wash. Year Book 58, 134 (1958).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Values for the x-ray lattice constant of the spinel $FeAl_2O_4$.

a_0 (Å)	Sample source	Reference	Ref. No.
8.146±0.004[a]	natural crystals, some Mg	Holgersson (1927)	16
8.135±0.002	synthetic powder	Clark, Ally, Badger (1931)	17
8.100±0.002	synthetic powder	Krause and Thiel (1931)	18
8.13	synthetic powder	Tazaki and Kuwabara (1952)	19
8.125±0.002	synthetic powder	Hoffmann and Fischer (1956)	20
8.152±0.001	synthetic powder, 1000°C	Atlas and Sumida (1958)	21
8.150±0.002	synthetic powder, 800°C	Turnock (1958)	22
8.152±0.002	synthetic powder, 1200°C	Roth (1963)	5
8.140±0.001	synthetic single crystals, 1200°C	Slack (present work)	...

[a] 8.155±0.004 Å for pure $FeAl_2O_4$ when corrected for slight amount of $MgAl_2O_4$.

gave an atomic ratio of aluminum to iron of 2.00±0.04. Even though the Al to Fe ratio is exactly 2.00, some of the Fe is probably in the Fe^{3+} state. Such oxidation of the Fe will increase the O to Fe ratio above the ideal value of 4. The chemical analysis of R67 gave this ratio as 4.1±0.1, indicating that some of the Fe was in the Fe^{3+} state. If the O to Fe ratio is actually 4.1, then 20% of the Fe occurs as Fe^{3+}. Except for the possibility that some of the Fe is in the Fe^{3+} state, the present synthetic crystals are close to being ideal $FeAl_2O_4$. A heat treatment of the as-grown crystals is needed in order to reduce the Fe^{3+} concentration to its lowest possible value. Crystal R75 was prepared by taking an as-grown sample of $FeAl_2O_4$, surrounding it with high-purity Al_2O_3 powder, and sealing it inside a capsule of high-purity iron. This capsule was then annealed at 1200°C for 21 h in an effort to bring the volume of the crystal to equilibrium with the reaction[12,21]:

$$FeAl_2O_4 \rightleftarrows Fe + Al_2O_3 + \tfrac{1}{2}O_2.$$

This equilibration produced a slight increase in a_0 to 8.140 ±0.002 Å and, as will be shown later, a noticeable decrease in the optical absorption, and a sharpening of the kink in the thermal conductivity curve at 10°K. Since the equilibration is a diffusion process, annealing times even longer than 21 h may be necessary to bring the bulk of the crystal to equilibrium with the oxygen atmosphere. A thin sample of $FeAl_2O_4$ which was given a longer anneal had a lattice constant of 8.144 Å, which is approaching the ideal value of 8.151 Å. The fact that the a_0 of the synthetic $FeAl_2O_4$ is less than 8.151 Å is attributed to the presence of Fe^{3+}.

Chemical Purity

The chemical purity of the various samples has been measured in order to determine what foreign elements were present. This is especially necessary since many of the samples were natural crystals. The impurity content of samples R42, R54, R56, and R62 has previously been published.[23] The major impurities in the other four samples are listed in Table III. These have been determined by spectrographic analysis in all cases except for the Fe and Cr of R97. These two impurities were determined by aqueous solution colorimetry to ±10%. These results show that Fe is the dominant impurity in R97. Crystal R97 was in fact chosen from a group of several natural spinel crystals just for this reason.

The synthetic $FeAl_2O_4$ is reasonably pure. Most of the impurities appear to have been introduced during the growing process since Johnson-Matthey "Specpure" oxides were used as the starting material. With more care purer samples could be made.

THERMAL CONDUCTIVITY

The physical properties that were measured were the lattice constants a_0 (see Table I), the thermal conductivity, the optical absorption, and the magnetic susceptibility. The interpretation of and the interrelationships between these parameters will be discussed.

Extensive measurements of the thermal conductivity K have been made between 3 and 300°K by techniques previously described.[24] The K versus T curves for R42, R54, R56, and R62 have already been published.[23] They are reproduced again in Fig. 2 along with the new data for the other crystals. The main features to notice in Fig. 2 are the monotonic decrease in K at any and all temperatures with the increasing content of Fe. The pure $FeAl_2O_4$ has the lowest K of all. This crystal is a rare example of a good single crystal with no (or very little) atomic disorder in which the K decreases continuously with decreasing T for all temperatures below 300°K. The other important feature to notice is the "kink" in several of the K versus T curves at 10°K.

TABLE III. Major metallic impurities in several spinel samples in units of 10^{18} atoms/cm^3.

Impurity	Sample R68	Sample R67, R75	Sample R97
Ca	8	<0.7	<5
Cr	<0.2	3	35
Fe	0.4	m[a]	360
Mg	m	0.3	m
Si	2	7	4
Ti	0.5	1	5
V	<0.4	0.3	8
Zn	<0.7	<0.8	20

[a] m = major constituent at a concentration of about 1.5×10^{22} cm^{-3}.

[23] G. A. Slack, Phys. Rev. 126, 427 (1962).
[24] G. A. Slack, Phys. Rev. 122, 1451 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

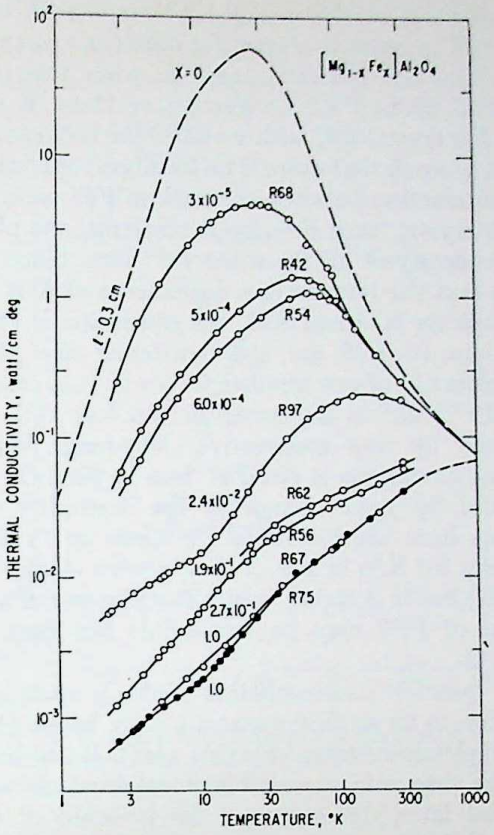


FIG. 2. The thermal conductivity K versus temperature for several mixed crystals of $Mg_{1-x}Fe_xAl_2O_4$. The value of x for each crystal is given. Note the continuous decrease in K with increasing x.

The heat transport in all of these crystals is believed to be caused by phonons. It does not appear that the magnetic moments in any of the crystals become well ordered between 3 and 300°K, so that a magnon heat transport is not at all likely. This lack of ordering will be discussed later. The phonons in the pure $MgAl_2O_4$ with $x=0$ are scattered by[23] the crystal boundaries, the isotopes, and the other phonons for $3°K \leq T \leq 300°K$. The relative strengths of these scattering processes vary with temperature, and their combined effect gives an upper bound to the K given by the dashed curve labeled $x=0$ in Fig. 2. This curve is computed for a specimen diameter of 0.3 cm, which is just the average crystallite size in the purest sample R68. The other crystals ranged in diameter from 0.2 to 0.4 cm, and in length from 0.3 to 1.2 cm. Thus, 0.3 cm is a reasonable average diameter for all of the crystals. Only the purest sample R68 comes close to the K limit imposed by boundary scattering at 3°K.

The failure of sample R68 to have a K as high as the dashed curve in Fig. 2 may be caused by some disorder in the Mg and Al distribution on the A and B sites that has been seen in other synthetic spinels.[4] It may also be caused by the presence of the various chemical impurities listed in Table III. The Fe and Ti are the dominant transition-metal impurities, and are the most likely source of any phonon scattering by the impurities.

In these crystals containing tetrahedral Fe^{2+} it is suggested in this paper that the isolated Fe^{2+} ions are very effective phonon scatterers. The scattering mechanism is believed to be the same as that postulated for Fe^{2+} in the tetrahedral sites in CdTe,[25] and is caused by phonon assisted transition between the several low-lying energy levels of the $3d^6$ configuration. Thus, the phonon scattering and hence the thermal resistivity $W=K^{-1}$ should depend on the Fe^{2+} concentration. Notice that in Fig. 2 for a crystal with $x=0.1$, i.e., somewhere between R97 and R62, the K would decrease monotonically with decreasing T for $T \leq 300°K$. This means that for $x>0.1$, the phonon scattering by the Fe^{2+} predominates over the normal phonon-phonon scattering for all $T \leq 300°K$. In order to characterize the Fe^{2+} scattering mechanism, we are interested in the K as a function of T and of x.

Values of x for the concentration of tetrahedral Fe^{2+} in the various crystals have been obtained from the strength of the optical absorption at 0.60 eV. The details are described later. These x values are given in Table I. The x values for R67 and R75 will change slightly with heat treatment and the residual concentration of Fe^{3+}, however, they are probably between 0.9 and 1.0. Figure 3 shows a plot of the thermal resistivity W versus x at 3, 10, 30, 100, and 300°K for the crystals studied. The straight lines in Fig. 3 are drawn to weight the points for R97 and R75 most heavily since in these two crystals Fe^{2+} is by far the most prevalent impurity. Crystals R62 and R56 have substantial concentrations of Fe^{3+}, as found in the optical absorption studies. Crystals R54 and R68 both have several other impurities whose concentrations are equal to or greater than that of the Fe. Figure 3 indicates, however, that Fe^{2+} is responsible for a major part of the W in these two crystals between 3 and 30°K. The lines in Fig. 3

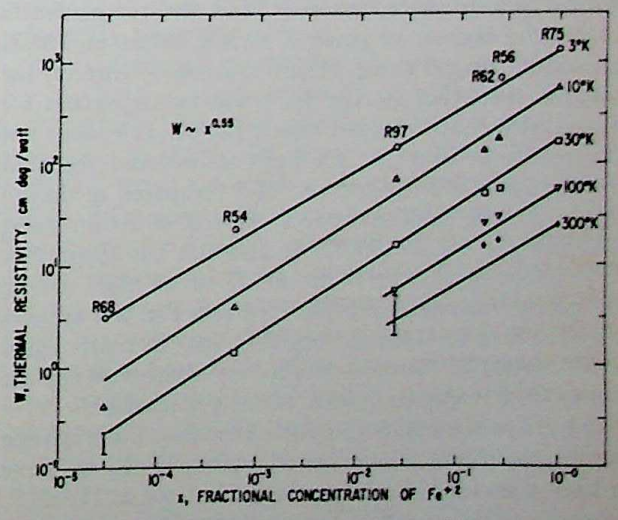


FIG. 3. The thermal resistivity W versus the Fe^{2+} concentration x for several temperatures. For any given temperature, w varies approximately as $x^{0.55}$.

[25] G. A. Slack and S. Galginaitis, Phys. Rev. 133, A253 (1964).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

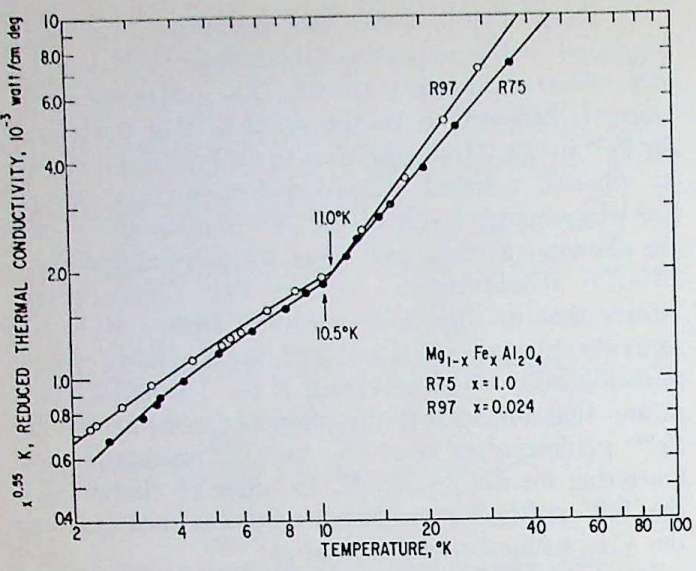


FIG. 4. The reduced thermal conductivity versus temperature in the region of 10°K. The similarity of the "kink" at 11°K for R75 and R97 is quite plain.

show that when the Fe^{2+} is the dominant phonon scatterer, W varies approximately as $x^{0.55\pm0.05}$, or K as $x^{-0.55}$ for all temperatures between 3 and 300°K. The K versus T curves in Fig. 2 show that in the temperature range where Fe^{2+} is the dominant phonon scatterer, K varies roughly as T^{1+}. A detailed calculation of the magnitude, temperature, and Fe^{2+} concentration dependence of K for the crystals shown in Fig. 2 could be made using Callaway's[26] approach if one had a good idea of the relaxation times for phonon scattering by the Fe^{2+} ions and by the normal phonon processes. However, neither is very well known at present. The scattering from the Fe^{2+}, $3d^6$ configuration in tetrahedral sites involves a series of five equally-spaced magnetic levels, and is similar to the case of Fe^{2+} in CdTe.[25] Some recent, preliminary measurements by the author on the optical absorption of crystal R97 at low temperatures places the highest of these 4 excited levels at 120°K above the ground state. If this estimate is correct, the phonon scattering by the Fe^{2+} will be important for temperatures up to a few times 120°K. It is thus not difficult to believe that the Fe^{2+} has reduced the K of $FeAl_2O_4$ at 300°K by a factor of 5 compared to that of $MgAl_2O_4$. The continuous decrease in K with decreasing temperature for $T<300$°K is also not too surprising. The effect of the Fe^{2+} on the K of $FeAl_2O_4$ should disappear only when $T\gg120$°K. From Fig. 2 it appears that the K of $FeAl_2O_4$ and $MgAl_2O_4$ may be nearly equal for $T>1000$°K. A more detailed calculation to explain the $x^{-0.5}$, T^{1+} variation of K is not yet feasible.

The curves for R75 and R97 show that there is some structure in the K versus T curves. Both of these exhibit a decided "kink" or change in slope at 11±1°K. In order to demonstrate the similarity between the two curves, reduced K values for both of them have been replotted side-by-side in Fig. 4. The reduced K is taken as $x^{0.55}$ K in order to correct for different x values. The "kink" for R97 comes at 11±1°K, while that for R75 comes at 10.5±1°K. An average of 11±1°K will be used. For crystal R97 with $x=0.024$ the Fe^{2+} concentration is so small that there is no likelihood of a magnetic A-A interaction between the various Fe^{2+} ions. Thus, in this crystal, each Fe^{2+} ion is scattering the phonons independently of all the other Fe^{2+} ions. Since Fig. 4 shows that the temperature dependence of K is nearly the same for R75 and R97, the conclusion is that the Fe^{2+} ions in R75 are also scattering the phonons independently of one another for $x=1$. It is concluded that the "kink" in the curves in Fig. 4 at 11°K is not produced by any cooperative, long-range magnetic ordering among the A site Fe^{2+} ions in $FeAl_2O_4$, but is produced by some change in the scattering of the phonons from the individual Fe^{2+} ions at 11°K. The K results for R56 in Fig. 3 show a very small "kink" at 11°K, but it is rather weak. The presence of a large amount of Fe^{3+} may be responsible for suppressing the effect.

The question as to whether $FeAl_2O_4$ undergoes a transition to an antiferromagnetic state below 11°K is not completely answered by this result. Roth[5] believes that the magnetic susceptibility data, which will be discussed later, demonstrates the presence of short-range antiferromagnetic ordering below 8 to 10°K. His neutron results, however, fail to show any long-range ordering at these temperatures. The present K results show that if a transition to an antiferromagnetic state does take place in $FeAl_2O_4$, such a transition does not affect the K. This would be surprising in view of the results[24,27,28] on MnO, CoO, MnF_2, and CoF_2; on[29] $CuCl_2\cdot2H_2O$ and $CoCl_2\cdot6H_2O$, and on[30] UO_2, where there was a noticeable break in the K versus T curve on passing through the Néel temperature. More work will be necessary before the 11°K transition in $FeAl_2O_4$ is understood.

As a check on the hypothesis that the Fe^{2+} in the $FeAl_2O_4$ is the cause of the anomalously low thermal conductivity, some measurements of K were made on a single crystal of $CoAl_2O_4$. This crystal R106 is a normal spinel[5] and was grown by E. M. Clausen of this laboratory using a flame-fusion technique. The K results for a crystal 0.83 cm×0.20 cm×0.18 cm annealed in air for 100 h at 1000°C are shown in Fig. 5 together with the curves for $MgAl_2O_4$ (R68) and $FeAl_2O_4$ (R75). The K results for the as-grown $CoAl_2O_4$ sample were identical to those for the annealed sample. Notice that the K values lie between the other two curves except at the

[26] J. Callaway, Phys. Rev. 122, 787 (1961).

[27] G. A. Slack and R. Newman, Phys. Rev. Letters 1, 359 (1958).

[28] G. A. Slack, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Academic Press Inc., New York, 1962), p. 630.

[29] R. H. Donaldson and D. T. Edmonds, Phys. Letters 2, 130 (1962).

[30] O. Bethoux, P. Thomas, and L. Weil, Compt. Rend. 253, 2043 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

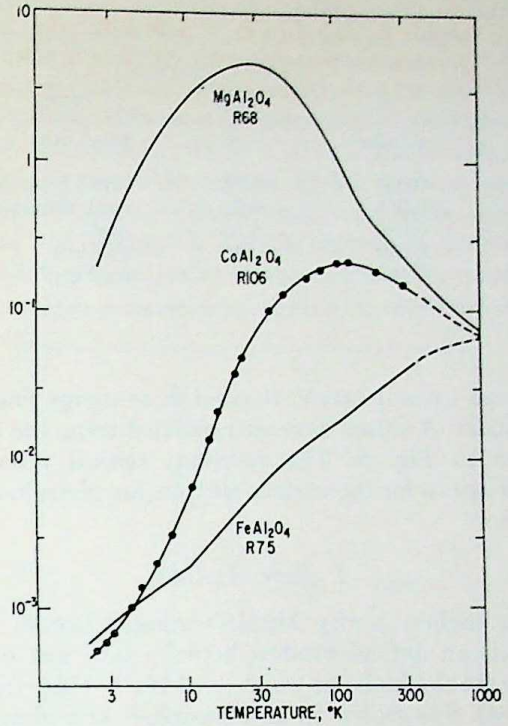


FIG. 5. The thermal conductivity of a single crystal of $CoAl_2O_4$ versus temperature. The results lie between those for $MgAl_2O_4$ and $FeAl_2O_4$, which are also shown.

very lowest temperatures. The K of $CoAl_2O_4$ does not decrease monotonically with decreasing T. All three crystals appear to have nearly identical K for $T \geq 1000$°K. The tetrahedral Co^{2+} has a $3d^7$ configuration, and does not possess[25] any low-lying energy levels in the 1 to 1000°K energy range. Thus it should not scatter phonons with anywhere near the same effectiveness as does the tetrahedral Fe^{2+}. This conclusion is demonstrated by the results in Fig. 5, at least for temperatures between 4 and 300°K. Solely on the basis of the magnetic scattering by the tetrahedral Co^{2+}, the K of $CoAl_2O_4$ should be nearly as high as that of $MgAl_2O_4$. The failure of the $CoAl_2O_4$ to exhibit such a high K may be caused by a partial inversion of the $CoAl_2O_4$, or by other crystal imperfections.

From the differences in K of the spinels $MgAl_2O_4$, $FeAl_2O_4$, and $CoAl_2O_4$, one concludes that the phonon scattering is very sensitive to the details of the magnetic nature of the A-site ions.

OPTICAL PROPERTIES

Optical absorption measurements on most of the samples listed in Table I and on a sample of commercial $MgO \cdot 3.5\ Al_2O_3$ spinel have been made at 300°K over the photon-energy range from 0.1 to 7.0 eV (12- to 0.18-μ wavelength). These measurements were made using three different recording spectrophotometers. A Perkin-Elmer[31] model-421 grating instrument was used from 0.1 to 0.5 eV, and a Beckman[32] DK-2A and a Cary[33] model-14 were used from 0.5 to 7.0 eV. These instruments operated well down to about 1% sample transmission. The sample thickness ranged from 5×10^{-1} to 3×10^{-3} cm. Samples thinner than 3×10^{-3} cm were difficult to prepare because they often broke during the polishing process. This combination of instruments and samples allowed one to measure values of the optical absorption coefficient α from 10^{-1} to 10^{+3} cm^{-1}. The coefficient α is defined by

$$I = I_0(1-R)^2 e^{-\alpha t}, \quad (1)$$

where I_0 = incident light intensity, I = transmitted light intensity, R = reflection loss for a single surface at normal incidence, t = sample thickness in cm.

The two main purposes of these measurements are to determine the site symmetry of the Fe^{2+} in the spinel crystals and to determine the concentration of the Fe^{2+}. The Fe^{2+} in the tetrahedral sites produces a strong absorption in the infrared over a broad band between 0.4 and 0.8 eV. The α-versus-photon energy $h\nu$ curves allow one to make some estimate of the Fe^{2+} and the Cr^{3+} concentration in the crystal.

Reflectivity

Some optical reflectivity measurements were made on $MgAl_2O_4$ and $FeAl_2O_4$ from 0.03 to 2.0 eV in order to identify the regions of strong lattice absorption and to measure the reflection loss R. These measurements were made at 300°K with a Perkin-Elmer[31] model-83 monochromator at nearly normal incidence. The strong lattice peaks were all found in the energy range between 0.03 and 0.11 eV. The samples used were a natural spinel crystal from Burma which was medium-pink in color, and a synthetic crystal of $FeAl_2O_4$ in the as-grown, unannealed condition. The surfaces of the crystals were polished so that they were optically flat before the reflectivity was measured.

The reflectivity results for $MgAl_2O_4$ are shown in Fig. 6. The main lattice absorption peaks are at 0.066 and 0.091 eV. The other smaller peaks that were found are listed in Table IV along with those found for $FeAl_2O_4$. The reflectivity values for $MgAl_2O_4$ in Table IV for normal incidence are reasonably accurate. The values for $FeAl_2O_4$ taken at about 30° from the normal are only approximate, but do indicate the relative strengths of the various peaks. Table V also gives the positions for the absorption maxima by transmission found by other authors[4,24,35] using powdered samples compressed in KBr pellets. The Greek letters serve to identify the various peaks. Even though the peaks in the reflectivity and absorption are not at

[31] Perkin-Elmer Corporation, Norwalk, Connecticut.

[32] Beckman Instruments Inc., Fullerton, California.

[33] Applied Physics Corporation, Monrovia, California.

[34] K. A. Wickersheim and R. A. Lefever, J. Opt. Soc. Am. 50, 831 (1960).

[35] S. Hafner, Z. Krist. 115, 331 (1961).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IV. Reflectivity maxima found for $MgAl_2O_4$ and $FeAl_2O_4$ single crystals.

Designation (a)	$MgAl_2O_4$ (natural) $h\nu$	$\%R$	Character	$FeAl_2O_4$ (synthetic) $h\nu$	$\%R$	Character
α	0.0375 eV	36	sharp peak	0.038 eV	~50	small peak
β	0.059	89	shoulder			
γ	0.066	100	broad peak	0.061	~70	broad peak
δ	0.072	75	shoulder	0.068	~50	weak shoulder
ϵ	0.091	100	sharp peak	0.086	~80	peak
ζ	0.104	50	weak shoulder	0.099	~40	shoulder
θ	0.176	small	see Fig. 7	0.18	low	see Fig. 7

[a] These Greek letters serve only to identify the various peaks.

identical energies, the energies are nearly equal. The strong reflectivity peaks are associated with the strong absorption ones. There is a close similarity between these lattice vibration peaks for $MgAl_2O_4$ and $FeAl_2O_4$ since they both have the spinel crystal structure, and nearly the same lattice constants. The Fe is a heavier atom than the Mg and the reflectivity peaks for $FeAl_2O_4$ occur at an energy about 5% less than those for $MgAl_2O_4$. In the energy range between 0.12 and 3.0 eV (and up to 4 eV actually) the reflectivity R of $MgAl_2O_4$ is quite small (2 to 8%), and there is no structure in the R versus $h\nu$ curve in Fig. 6. The same general statement can be made for $FeAl_2O_4$. Thus, any structure found in the optical transmission curves of the single crystals in this energy range is caused by variations in the absorption coefficient α and the reflection correction is small, see Eq. (1).

Absorption Coefficient

The % transmission curves that were taken were corrected for reflection losses using R values calculated from the refractive index[36,37] n for $MgAl_2O_4$ in the range between 1.6 and 3.0 eV. Beyond these energy limits the calculated R values were extrapolated using the results shown in Fig. 5. The resultant optical absorption coefficients α for the various samples are given in Figs. 7 and 8.

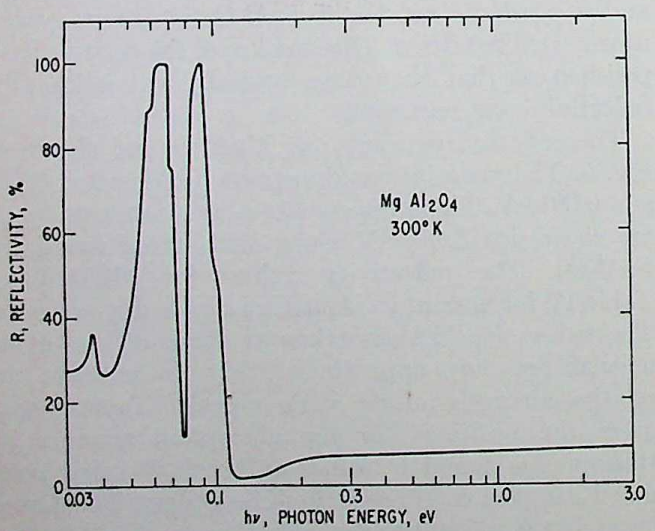


FIG. 6. The optical reflectivity R versus photon energy. These results for a natural single crystal of $MgAl_2O_4$ at 300°K were taken at nearly normal incidence.

1. Pure $MgAl_2O_4$

The highest purity $MgAl_2O_4$ sample R68 in Fig. 7 exhibits an optical window between 0.23 and 6.0 eV. These are the limits at which $\alpha = 1$ cm^{-1}. This synthetic $MgAl_2O_4$ was grown[6] in an atmosphere of hydrogen by diffusing MgO into Al_2O_3. It does not show any signs of the OH or H_2O band at 0.4 eV (3 μ) that has been seen[34,38] in synthetic spinels grown by the flame-fusion process. As a check on the presence of this band a synthetic crystal of $MgO+3.5\,Al_2O_3$ from Linde[39] (crystal R53 of Ref. 23) was measured in this region. It showed a double-peaked band with peaks at 0.415 and 0.436 eV with absorption coefficients of 6 cm^{-1} and 3 cm^{-1}, respectively. For contrast, the R68 crystal had an $\alpha < 0.05$ cm^{-1} in this same energy region. Thus, the 0.4-eV band is not an intrinsic feature of $MgAl_2O_4$, but shows up only in the flame-fusion grown crystals. It has not been seen in any of the natural $MgAl_2O_4$ crystals. It is necessary to understand the source of this band, since the absorption peak for Fe^{2+} ions in the A sites occurs at about this same energy.

TABLE V. Absorption maxima for $MgAl_2O_4$ and $FeAl_2O_4$ powders from the literature.

Designation (a)	$MgAl_2O_4$ (b)	(c)	(d)	$FeAl_2O_4$ (e)
γ	0.0645 eV	0.067	0.065	0.063
δ	0.0716			0.070
ϵ	0.0849	0.086	0.085	0.083
ζ	0.0932	0.092	0.097	
η			0.124	

[a] These Greek letters serve only to identify the various peaks.
[b] Natural $MgAl_2O_4$, Ref. 4.
[c] Synthetic $MgAl_2O_4$, Ref. 35.
[d] Synthetic $MgAl_2O_4$, Ref. 34.
[e] Synthetic $FeAl_2O_4$, Ref. 4.

[36] F. Rinne, Neues Jahrb. Mineral. Geol., Beilageband 58A, 43 (1928).
[37] K. Schlossmacher, Z. Krist. 72, 447 (1930).
[38] G. Calingert, S. D. Heron, R. Stair, Soc. Autom. Engin. J. 39, 448T (1936).
[39] Linde Air Products Company, New York, New York.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2. *Tetrahedral* Fe^{2+}

The main feature in all of the α versus $h\nu$ curves in Figs. 7 and 8 is the very strong absorption band between 0.4 and 0.8 eV. This band is assigned to the tetrahedral Fe^{2+}, $3d^6$ transition $^5E(^5D) \rightarrow {}^5T_2(^5D)$. This same transition has been seen in several II–VI semiconducting compounds doped with Fe. Probably the earliest report of it was by Coblentz[40] in natural crystals of zincblende (cubic ZnS), which almost always contain Fe, where he found a broad single absorption peak at 0.41 eV. Low and Weger[41] have also found this peak in natural zincblende, Pappalardo and Dietz[42] have seen it in Fe-doped synthetic CdS, and the present author[43] has seen it in Fe-doped synthetic CdTe. The energies of the two peaks in this double peak are listed in Table VI. In some cases only a single, broad peak is seen. This absorption peak shifts toward higher energies as the Fe-anion distance decreases, see Table VI.

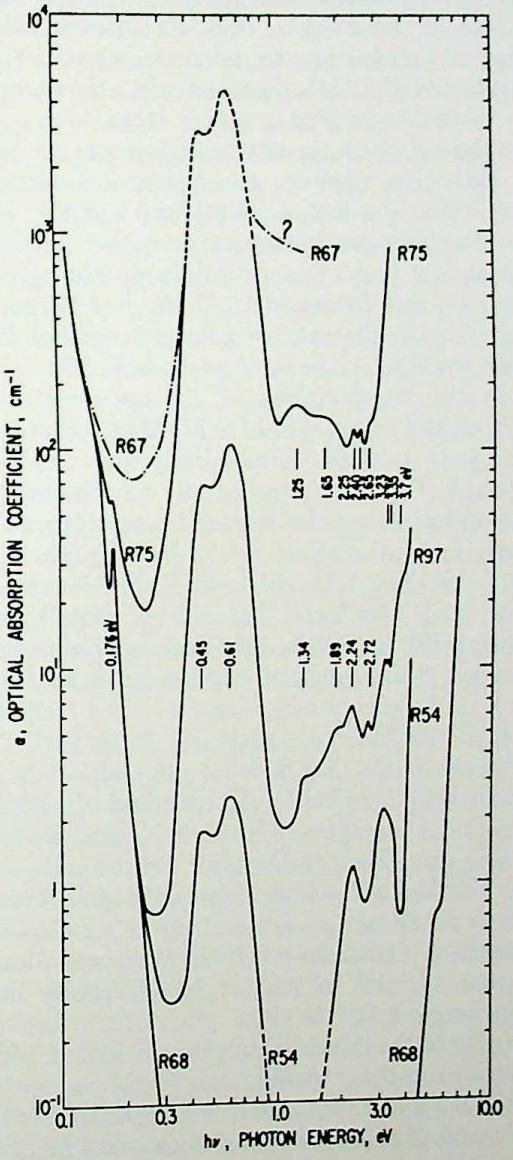


FIG. 7. The optical absorption coefficient α versus the photon energy $h\nu$ for several different single-crystal spinels. Corrections for reflection losses have been made. The energies of various absorption peaks have been labeled in electron volts. The dashed curve for R75 indicates that the absorption between 0.4 and 0.8 eV was too large to be measured. However, the estimated behavior is shown. Similarly the ? for R67 indicates that above 0.8 eV, α was too large to be measured. The main absorption peaks at 0.45 and 0.61 eV are caused by tetrahedral Fe^{2+}.

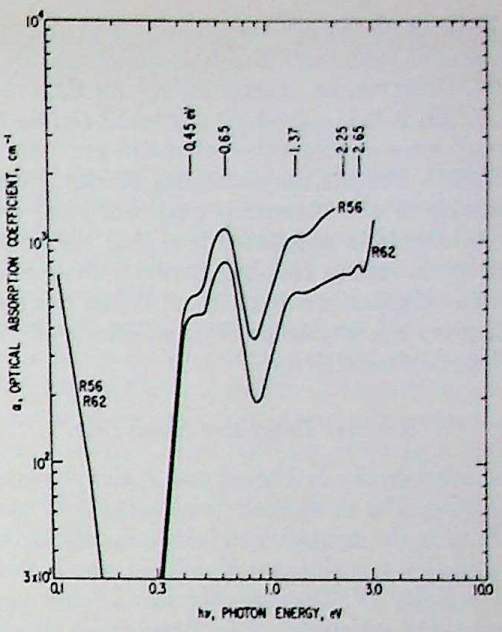


FIG. 8. The optical absorption coefficient α versus the photon energy $h\nu$ for two pleonaste spinels which contain large amounts of both Fe^{2+} and Fe^{3+}. Corrections for reflection losses have been made. The main peaks are labeled in electron volts.

The strength of the electrostatic crystalline field at the Fe^{2+} ion produced by the other surrounding ions is given[44] by the parameter Dq. The energy $h\nu$, at the peak in the α versus $h\nu$ curve is approximately equal to 10 Dq.

TABLE VI. Peak energies for the double-peak, infrared optical absorption band of tetrahedral Fe^{2+} at 300°K.

Crystal	Peaks $h\nu$, eV	Reference	Fe-anion distance, Å
$MgAl_2O_4$, cubic	0.45, 0.61	present work	1.81
ZnS, cubic	0.41	Coblentz[b]	2.34
ZnS, cubic	0.35, 0.50	Low and Weger[c]	2.34
CdS, hex.	0.306, 0.387	Pappalardo and Dietz[d]	2.53
CdTe, cubic	0.34[a]	Slack[e]	2.81

[a] Splits into two peaks when crystal is cooled below 100°K.
[b] See Ref. 40.
[c] See Ref. 41.
[d] See Ref. 42.
[e] See Ref. 43.

[40] W. W. Coblentz, No. 97, Carnegie Institution of Washington, D. C., 1906 (unpublished), Part VI, p. 57.

[41] W. Low and M. Weger, Phys. Rev. 118, 1119, 1130 (1960).
[42] R. Pappalardo and R. E. Dietz, Phys. Rev. 123, 1188 (1961).
[43] G. A. Slack (unpublished).
[44] D. S. McClure, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1959), Vol. 9, p. 399.

CC-0. In Public Domain. G

Digitized by Arya Samaj Foundation Chennai and eGangotri

The breadth of the absorption band and the presence of the double peak make the determination of Dq rather inexact. However, an average energy for the two peaks in $MgAl_2O_4$ is 0.53 eV = 4300 cm^{-1} = 10 Dq. So Dq has the not unreasonable value of +430 cm^{-1} for tetrahedral Fe^{2+}. Further understanding of what causes the broadening of the absorption peak will yield a better value of Dq. It is postulated here that the breadth is caused by a strong coupling between these magnetic levels and the lattice vibrations. It is just this coupling which gives rise to the pronounced effects of Fe^{2+} on the thermal conductivity.

3. *Other Transition Metal Ions*

The other peaks in Figs. 7 and 8 also require some explanation. The small, sharp peak at 0.176 eV has been seen in all of the crystals, and is independent of doping. It is probably a lattice vibration band, and is therefore listed in Table IV. There are numerous peaks between 1.0 and 4.0 eV which vary from sample to sample. Since pure $MgAl_2O_4$, sample R68 has no absorption bands between 0.23 and 6.0 eV, these bands are caused by various transition-metal impurities. So far none of these other bands in Figs. 7 and 8 has been definitely associated with transitions in the tetrahedral Fe^{2+} ion. They all appear to be caused either by octahedral Cr^{3+} in the natural crystals (R54 and R97), by Fe^{3+}, or by octahedral Fe^{2+} in the other crystals. The various Cr^{3+} peaks in R54 and R97 are dealt with in the Appendix. The conclusion is that except for the 1.34-eV peak in R97 all the other peaks above 1.0 eV in these two crystals are due to Cr^{3+}. Crystal R54 shows the two peaks at 2.24 and 3.23 eV characteristic of octahedral Cr^{3+}. In R97 the simple octahedral Cr^{3+} spectrum has been modified by the presence of the much larger concentration of tetrahedral Fe^{2+}.

4. *Other Fe Ions*

Let us now consider the possibilities involved in the absorption spectra of Fe^{3+}, both octahedral and tetrahedral, and of octahedral Fe^{2+}. If $FeAl_2O_4$ were a perfectly stoichiometric, normal spinel, then only tetrahedral Fe^{2+} would exist. The x-ray data have shown[1,2] that $FeAl_2O_4$ is a normal spinel. The recent work of Schmalzried[45] on other aluminates has shown that $CoAl_2O_4$ is nominally a normal spinel, but when equilibrated at 1000°C is partially inverse. In view of the similarity of the octahedral preference energies of $CoAl_2O_4$ and $FeAl_2O_4$ calculated by Miller,[46] $FeAl_2O_4$ can be expected to be slightly inverse in the 1000 to 1500°C range. From neutron scattering data, Roth[5] has found that in his powder sample of synthetic $FeAl_2O_4$, annealed at 1200°C, about 15% of the iron was in the octahedral sites. Thus there will probably be some small amount (i.e., 1 to 20%) of octahedral Fe^{2+} in the present single crystals annealed at 1200°C. The presence of Fe^{3+} in the $FeAl_2O_4$ is less certain. It is well known[14] that the iron in wüstite or "FeO" cannot be reduced to the point where all of the iron is divalent. Thus, one might expect some Fe^{3+} in the $FeAl_2O_4$ even when it is at the low-oxygen pressure dissociation limit of decomposing into Fe, Al_2O_3, and O_2. From Miller's[46] calculations one would predict that this Fe^{3+} would be on the tetrahedral sites. Notice that any octahedral Fe^{3+} that is present requires only that one of the many tetrahedral Fe^{2+} ions in the $FeAl_2O_4$ transfers an electron to the octahedral Fe^{3+} in order to yield an octahedral Fe^{2+} and a tetrahedral Fe^{3+} ion. For example, this exchange is quite favorable in Fe_3O_4, which is an inverse spinel where the Fe^{3+} is on the tetrahedral sites in competition with Fe^{2+}. The opposite conclusion of Hoffman and Fischer[20] that the first small additions of Fe^{3+} go into the octahedral sites is suspect since their lattice constants at low Fe_3O_4 concentrations do not agree with the later ones of Atlas and Sumida[21] or of Turnock.[22] It seems probable that at low concentrations of Fe^{3+} in $FeAl_2O_4$ the Fe^{3+} occurs on the tetrahedral sites. Therefore, in the interpretation of the optical absorption spectra of the mixed $FeAl_2O_4$–$MgAl_2O_4$ crystals, it is necessary to consider tetrahedrally and octahedrally coordinated Fe^{2+}, and tetrahedrally coordinated Fe^{3+}.

The optical absorption caused by octahedral Fe^{2+} is believed to be the cause of the broad peaks seen in Fig. 6 for R75 at 1.23 and 1.65 eV. Previous work by Holmes and McClure[47] has shown that hydrated, octahedral Fe^{2+} in $FeSO_4 \cdot 7H_2O$ crystals possesses only two rather weak, broad absorption bands at 1.04 and 1.34 eV in the energy range from 0.75 to 4.7 eV, with a combined oscillator strength of $f \simeq 4 \times 10^{-5}$. These authors also found that the absorption of Fe^{2+} in water solution was very similar. The method of calculating the oscillator strengths f from the optical absorption curves is that given by Dexter.[48] Octahedrally coordinated Fe^{2+} is a well-known colorant in glass,[49] where it produces an absorption band at about 1.2 eV (1 μ) in the infrared. Octahedral Fe^{2+} in MgO was found by Low and Weger[41] to have a broad, poorly defined absorption peak at 1.24 eV.

A study of the optical absorption of Fe^{2+} in MgO by the present author[43] shows, that in the energy range between 0.5 and 3.0 eV, there is a broad, rather weak, double-peaked absorption band with peaks at 1.21 and 1.46 eV. The derived oscillator strength is $f \simeq 6 \times 10^{-5}$ for each peak,[6] close to the value for the hydrated Fe^{2+}. The energies for these absorption bands of Fe^{2+} in MgO and $FeSO_4 \cdot 7H_2O$ are in reasonable agreement with the

[45] H. Schmalzried, Z. Physik. Chem. (Frankfurt) 28, 203 (1961).
[46] A. Miller, J. Appl. Phys. 30, 245 (1959).
[47] O. Holmes and D. S. McClure, J. Chem. Phys. 26, 1686 (1957).
[48] D. L. Dexter, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1958), Vol. 6, p. 353, 371.
[49] W. A. Weyl, *Coloured Glasses* (Society of Glass Technology, Sheffield, 1951), Chap. 7.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

1.23- and 1.65-eV bands seen in $FeAl_2O_4$ in Fig. 6. This is the reason for the assignment of these two bands to octahedral Fe^{2+}. If one makes the assumption that 15% of all the Fe^{2+} in the R75 sample of $FeAl_2O_4$ is on the octahedral sites, and is producing the 1.23-eV band, Fig. 6, the calculated oscillator strength for this 0.7-eV wide band is $f \simeq 3\times10^{-4}$. This is five times higher than that found for MgO. Since the actual concentration of octahedral Fe^{2+} in R75 may be higher or lower than the 15% found by Roth, the f value is only approximate. If anything, the octahedral Fe^{2+} concentration in R75 is probably less than 15% of the total, thus making $f > 3\times10^{-4}$. The optical transition responsible for this broad, double-peaked band is believed to be the lowest energy one possible,[44] that is ${}^5T_2({}^5D) \rightarrow {}^5E({}^5D)$. The intensity of this optical transition and hence its f value is enhanced by distortion of the local cubic symmetry around the Fe^{2+} ion. The f value is small in highly symmetric environments such as at a Mg site in MgO. In $FeAl_2O_4$ the oxygen ions surrounding the octahedral site form a distorted octahedron. The site symmetry may be further reduced by the presence or absence of an Al^{3+} ion in one or more of the surrounding tetrahedral sites in the real crystal. These sites would, of course, be occupied only by Fe^{2+} in the pure, normal spinel. These effects will all serve to make the oscillator strength f of octahedral Fe^{2+} in $FeAl_2O_4$ greater than it is in MgO. Therefore, the optical absorption data is not yet of much help in determining the absolute concentration of octahedral Fe^{2+}, and hence the fractional inversion in $FeAl_2O_4$. However, it seems clear that the 1.23- and 1.65-eV bands are caused by octahedral Fe^{2+}, and may be useful in determining changes in the fractional inversion.

The other absorption peaks seen for R56 and R62 in Fig. 7 above 1 eV and for R75 in Fig. 6 above 2 eV are probably associated with the Fe^{3+} ions. Just what the assignments should be are rather uncertain. The lowest energy peak for Fe^{3+} in octahedral coordination with oxygen in α-Fe_2O_3 has been found by Bailey[50] at 1.49 eV. A similar peak at the same energy in $Y_3Fe_5O_{12}$(YIG) has been found by Dillon.[51] This is probably the ${}^6A_1({}^6S) \rightarrow {}^4T_1({}^4G)$ transition. The first transition for tetrahedral Fe^{3+} should be at an even higher energy since Fe^{3+} has a ${}^6S$ ground state and the crystalline field at a tetrahedral site is about half of that at an octahedral site. Therefore, the optical absorption characteristic of tetrahedral Fe^{3+} should occur for $h\nu > 1.5$ eV in these oxide spinels. The optical absorption of Fe^{3+} tetrahedrally coordinated by Cl^{-1} ions has been studied by Friedman.[52] He found a number of peaks starting at 1.7 eV and extending toward higher energies. It seems reasonable to suppose that the 2.25-, 2.40,- and 2.65-eV peaks in R56, R62, and R75 may be caused by tetrahedral Fe^{2+}. More work is clearly necessary before such an assignment is convincing.

5. Charge Transfer

The last feature of Figs. 7 and 8 that remains unexplained is the hump in the curves for samples R56, R62, and R97 at 1.35 eV. Samples R56 and R62 have fairly large concentrations of Fe^{3+} ions in them, as will be shown later. These Fe^{3+} ions are probably on tetrahedral sites, as are the more numerous Fe^{2+} ions. When the same cation in two different charge states exists on the same lattice site in a crystal, it is relatively easy for a photon to cause an electron transfer between the two differently charged ions. Such a process is believed to be the source of a rather broad, structureless absorption in NiO,[53] $Y_3Fe_5O_{12}$,[54] and Fe_3O_4.[55] It is suggested here that this charge transfer absorption commences at about 1.35 eV in the present spinel crystals, and accounts for the hump in the curves. It would then be responsible for the dull, yellowish-brown appearance of the pleonaste spinels, R56 and R62, by transmitted light. See Table I.

CONCENTRATION OF TETRAHEDRAL Fe^{2+}

The shape of all of the absorption curves in Figs. 7 and 8 between 0.3 and 0.9 eV is nearly the same. The absorption in this portion of the curve is caused by the presence of tetrahedral Fe^{2+}. Since the shapes are all similar, the value of α at, say, the 0.61-eV peak should be proportional to the concentration of tetrahedral Fe^{2+}. The Fe concentration in R97 has been carefully measured as $3.6\pm0.4\times10^{20}$ atoms/cm^3. Since there is no sign of any optical absorption due to octahedral Fe^{2+}, and since the 1.34-eV peak, which is the only peak that might be associated with the presence of Fe^{3+}, is very weak, it is concluded that tetrahedral Fe^{2+} is by far the dominant species of Fe in crystal R97. The peak absorption cross section σ_p is defined by

$$\sigma_p = \alpha_p N^{-1}, \qquad (2)$$

where α_p = absorption coefficient at the peak, and N = concentration of the absorbing ions. The computed value is $\sigma_p = 2.9\times10^{-19}$ cm^2 for the 0.6-eV peak in R97. From Fig. 7 the full width of the 0.6-eV peak at $\alpha = 0.5\alpha_p$ is 0.22 eV. This width and the value of σ_p of 2.9×10^{-19} cm^2 yields[48] an oscillator strength of $f = 4\times10^{-4}$. The 0.45-eV peak gives a somewhat smaller calculated value of $f = 1\times10^{-4}$.

The concentration of tetrahedral Fe^{2+} in the other crystals in Figs. 7 and 8 has been calculated on the assumption that tetrahedral Fe^{2+} concentration is proportional to α_p at 0.6 eV. The results are given in Table VII along with the total concentration of Fe

[50] P. C. Bailey, J. Appl. Phys. 31, 39S (1960).
[51] J. F. Dillon, J. Phys. Radium 20, 374 (1959).
[52] H. F. Friedman, J. Am. Chem. Soc. 74, 5 (1952).
[53] R. Newman and R. M. Chrenko, Phys. Rev. 114, 1507 (1959).
[54] R. A. Lefever and A. B. Chase, J. Chem. Phys. 32, 1575 (1960).
[55] W. E. Engeler (private communication).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VII. A comparison of the concentration of tetrahedral Fe^{2+} ions (optical) with the concentration of total Fe (chemical).[a]

Sample	Conc. tetra. Fe^{2+} atoms/cm³	Conc. Fe total atoms/cm³	x
R68	$<5\times10^{17}$	4×10^{17} *s*	3×10^{-5}
R54	0.9×10^{19}	2×10^{19} *s*	6.0×10^{-4}
R97	3.6×10^{20} *a*	3.6×10^{20} *w*	2.4×10^{-2}
R62	2.7×10^{21}	5.0×10^{21} *w*	1.9×10^{-1}
R56	4.0×10^{21}	6.2×10^{21} *w*	2.7×10^{-1}
R67, R75		1.5×10^{22} *w*	~1

[a] *a* = assumed to be identical to concentration of total iron. *s* = determined from an emission spectrograph to within a factor of two. *w* = determined from a wet chemical analysis to ±10%. *x* = fraction of tetrahedral sites occupied by Fe^{2+}.

determined chemically. For the $FeAl_2O_4$ crystal the absorption was too large to be measured, and the expected curve, shown dashed in Fig. 7 between 0.4 and 0.8 eV, has been estimated from the chemical determination. The agreement with the known portion of the curve is quite good, showing that the proportionality assumption is not too bad. In Table VII the agreement between the optical and chemical values for R68 and R54 is again satisfactory. However, for samples R62 and R56 the chemical concentration is higher than the concentration of tetrahedral Fe^{2+}. Judging from the chemical analyses of these samples in Ref. 23 and the optical absorption curves in Fig. 8, the remaining 40% of the Fe probably occurs as Fe^{3+}.

MAGNETIC SUSCEPTIBILITY

The magnetic sysceptibility χ of an as-grown single crystal of $FeAl_2O_4$ quite similar to sample R67 was measured from 1.9 to 290°K. The measurements were made by Kouvel and Hartelius with an apparatus[56] described previously. The susceptibility was independent of applied field for values of the magnetic field H, between 1.5×10^5 to 8×10^5 A/m. (2000 to 10 000 Oe.) The susceptibility in rationalized mks units is shown in Fig. 9. The susceptibility is defined as the ratio of the induced magnetic moment per unit volume (A/m) to the applied magnetic field (A/m), and is dimensionless. These χ values can, if desired, be converted to emu/cm³ by dividing by 4π. The x-ray density of $FeAl_2O_4$ is 4.26 g/cm³. There is a very sharp peak in the χ versus T curve at 9.5±0.5°K. The χ^{-1} curve for the single crystal above 9.5°K is also shown in Fig. 9 compared with the recent data of Roth[5] and Lotgering[57] on powder samples. The present results lie between these two previous curves, and indicate that the χ is somewhat dependent on the exact nature of the sample. Small changes in the fractional inversion and the Fe^{3+} content are probably responsible for the variability in χ above 10°K. They are presumably responsible for slight shifts in the peak temperature. Roth[5] found a peak at 8°K. These measurements have not yet been made on annealed single crystals similar to R75.

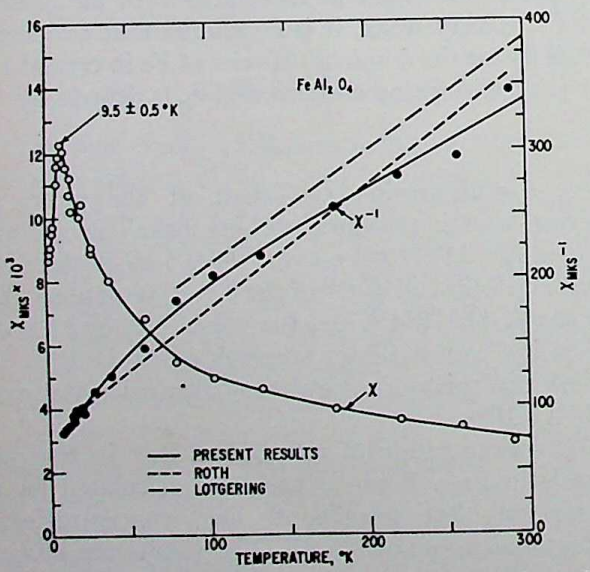


FIG. 9. The magnetic susceptibility χ and its reciprocal χ^{-1} versus temperature for a single crystal of $FeAl_2O_4$ similar to R67. The χ^{-1} curves obtained by other authors for powder samples are also shown.

The 9.5°K temperature for the peak in the χ versus T curve closely matches the "kink" in the K versus T curve at 11±1°K. It would seem that these two features are closely related. Since the "kink" in the K curve appears to be independent of the Fe^{2+} concentration, it is postulated here that the position of the peak may also be nearly independent of the Fe^{2+} concentration. Thus the behavior of the χ versus T curve may be a function of the individual Fe^{2+} ions, and may not be associated with a long-range ordering of the magnetic moment of the Fe^{2+}. A localized Jahn-Teller distortion of the Fe^{2+} ions at 11°K could, perhaps, account for much of the observed behavior. It is known that the normal spinel $FeCr_2O_4$ undergoes[58–60] such a distortion at 135°K, which is rather higher than its Curie temperature[58,61] of about 85°K. Even though Roth[5] found no evidence for a crystallographic distortion in the cubic $FeAl_2O_4$ at 4°K, it is possible that there is a small, local distortion around each Fe^{2+} site. This local distortion, however, may not be large enough to produce a long-range, cooperative, crystallographic distortion. The neutron scattering results of Roth also failed to show any evidence of long-range ordering in $FeAl_2O_4$ at 4°K. This rather unusual behavior of tetrahedral Fe^{2+} may be associated with the four low-lying energy levels, the closest of which is about 30°K above the singlet ground state. Since the ground state is a singlet with no orbital or spin degeneracy, the Fe^{2+} should have no magnetic moment in zero magnetic fields at temperatures well below 30°K. Thus, one might

[56] J. S. Kouvel, C. D. Graham, Jr., and J. J. Becker, J. Appl. Phys. **29**, 518 (1958).

[57] F. K. Lotgering, Phys. Chem. Solids **23**, 1153 (1962).
[58] C. H. Shomate, Ind. Eng. Chem. **36**, 910 (1944).
[59] E. Whipple and A. Wold, J. Inorg. Nucl. Chem. **24**, 23 (1962).
[60] A. Wold, R. J. Arnott, E. Whipple, and J. B. Goodenough, J. Appl. Phys. **34**, 1085 (1963).
[61] F. K. Lotgering, Philips Res. Rept. **11**, 190 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

expect that signs of a magnetic ordering at these temperatures would not occur. It would appear desirable to make χ measurements on mixed crystals of $FeAl_2O_4$–$MgAl_2O_4$ in order to test some of these speculations. This has not yet been done.

CONCLUSIONS

(1) A method of growing good single crystals of $FeAl_2O_4$ has been devised. A post-growth heat treatment in the appropriate atmosphere permits some control of the concentration of the residual Fe^{3+} and the fractional inversion.

(2) The thermal conductivity K of synthetic crystals of $FeAl_2O_4$ and $MgAl_2O_4$ and natural crystals of $Mg_{1-x}Fe_xAl_2O_4$ has been measured from 3 to 300°K. At all temperatures, the K decreases with increasing Fe^{2+} concentration x for $0 \leq x \leq 1$. For $x=1$, the K exhibits the unusual behavior of decreasing continuously with decreasing temperature for $T \leq 300$°K. In all of the crystals the heat is transported by phonons.

(3) The phonon scattering from the tetrahedral Fe^{2+} is quite strong. This scattering produces a change in the slope of the K versus T curve at 11 ± 1°K, independent of the Fe^{2+} concentration. These results suggest that the anomaly at 11°K is associated with the individual Fe^{2+} ions, and is not caused by the occurrence of an antiferromagnetic state.

(4) The magnetic susceptibility χ of a single crystal of $FeAl_2O_4$ has been measured from 1.9 to 290°K. The χ versus T curve shows a peak at 9.5 ± 0.5°K, in reasonable agreement with the 11°K peak found from the K measurements. It is suggested that this behavior may be characteristic of tetrahedral Fe^{2+}, and is not to be associated with the onset of an antiferromagnetic state.

(5) Thermal conductivity measurements on $CoAl_2O_4$ show that tetrahedral Co^{2+} is not as effective a scatterer of phonons as is tetrahedral Fe^{2+} in the range from 4 to 300°K.

(6) The optical absorption coefficients of crystals of $Mg_{1-x}Fe_xAl_2O_4$ for $0 \leq x \leq 1$ have been measured at 300°K for photon energies between 0.1 and 7.0 eV. The tetrahedral Fe^{2+} produces a broad, double-peaked absorption band between 0.4 and 0.6 eV. This absorption band has been used to measure the value of x.

Additional optical absorption bands have been seen which are due to octahedral Fe^{2+} and octahedral Cr^{3+}.

ACKNOWLEDGMENTS

I wish to thank J. H. McTaggart for his aid in growing the crystals and in making many of the measurements, R. M. Chrenko for his assistance with the optical absorption and reflectivity measurements, and Dr. J. S. Kouvel and C. C. Hartelius for measuring the magnetic susceptibility. I also thank E. M. Clausen for the $CoAl_2O_4$ crystal. My thanks are extended to W. Roth and F. S. Ham for several fruitful discussions concerning the nature of $FeAl_2O_4$.

APPENDIX: THE OPTICAL ABSORPTION OF Cr^{3+} IN $MgAl_2O_4$ SPINELS

Almost all natural crystals of $MgAl_2O_4$ contain traces of Cr. The chemical analyses in Table III and Ref. 23 show that Cr is indeed present in the natural crystals R54, R56, and R97. The Cr^{3+} ion has a very strong preference[46] for the octahedral sites in the spinel, and thus occurs only in these sites in $MgAl_2O_4$. The optical absorption of Cr^{3+} in spinel has been studied by many people,[62–71] and at low concentrations in natural crystals of reasonably pure $MgAl_2O_4$ it gives rise to two broad, intense absorption bands at 2.25 and 3.21 eV. Crystal R54 in Fig. 7 shows these two bands with the peaks at 2.30 and 3.18 eV, in good agreement with the previous literature. The emission spectrographic analysis of this sample showed a Cr content of 2×10^{19} cm^{-3} to within a factor of, say, 2. The calculated oscillator strengths[48] are, to within a factor of 2, $f = 2 \times 10^{-4}$ and 5×10^{-4} for the 2.30- and 3.18-eV peaks, respectively. If the synthetic crystal containing a nominal 0.5 weight % Cr studied by Schlossmacher is considered, the derived f value for his 2.21-eV peak is $f = 0.9 \times 10^{-4}$ to within a factor of 2. Thus the assignment of the 2.30-eV peak in R54 to Cr^{3+} is reasonable in terms of both the photon energy and the magnitude of the absorption coefficient. The optical absorption from a known amount of Cr^{3+} in α-Al_2O_3 has been measured by Maiman *et al.*[72] Their results for the 2.23- and 3.02-eV absorption bands averaged over the three directions in the crystal give f values of approximately 2×10^{-4} and 6×10^{-4}, respectively, for the two bands. These results for α-Al_2O_3 are in very good agreement with those for $MgAl_2O_4$. Of the other various references already cited for Cr^{3+} in $MgAl_2O_4$, none have enough data to enable a calculation of the oscillator strengths to be made.

Let us now consider the various bands found in R97 between 1.5 and 4.0 eV. The 2.24-eV band and the double-peaked band at 3.23 and 3.35 eV are believed to be the two Cr^{3+} bands that appear prominently in R54. The chemical analysis of R97 gave a Cr content of $3.5 \pm 0.3 \times 10^{19}$ cm^{-3}. From this value the calculated oscillator strengths are $f = 8 \times 10^{-4}$ for the 2.24 eV and $f = 16 \times 10^{-4}$ for the "3.29-eV" band. These f values are about 3 times larger than before. These two bands are

[62] O. Weigel, Neues Jahrb. Mineral. Geol., Beilageband 48, 274 (1923).
[63] O. Weigel and H. Ufer, Neues Jahrb. Mineral. Geol., Beilageband 57A, 397 (1928).
[64] K. Schlossmacher, Z. Krist. 72, 447 (1930).
[65] K. Schlossmacher, Z. Krist. 75, 399 (1930).
[66] K. Schlossmacher and I. Meyer, Z. Krist. 76A, 377 (1931).
[67] O. Deutschbein, Ann. Phys. 14, 712, 729 (1932).
[68] E. Kolbe, Neues Jahrb. Mineral. Geol., Beilageband 69A, 183 (1934).
[69] O. Schmitz-Dumont and D. Reinen, Z. Electrochem. 63, 978 (1959).
[70] R. A. Ford and O. F. Hill, Spectrochim. Acta 16, 1318 (1960).
[71] A. Neuhaus, Z. Krist. 113, 195 (1960).
[72] T. H. Maiman, R. H. Hoskins, I. J. D'Haenens, C. K. Asawa, and V. Evtuhov, Phys. Rev. 123, 1151 (1961).

CC-0. In Public Domain.

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE VIII. Comparison of the optical absorption transitions of Cr^{3+} in $MgAl_2O_4$ with those of $[Cr(H_2O)_6]^{3+}$.

Transition	Energy, $h\nu$, in eV R54	R97	$[Cr(H_2O)_6]^{3+}$
$^4A_2(^4F) \rightarrow {}^2E(^2G)$		1.89	1.85
$\rightarrow {}^2T_1(^2G)$			1.87
$\rightarrow {}^4T_2(^4F)$	2.30	2.24	2.23
$\rightarrow {}^2T_2(^2G)$		2.72	2.60
$\rightarrow {}^4T_1(^4F)$	3.18	3.23, 3.35	3.05
$\rightarrow {}^2A_1(^2G)$		3.7	
$\rightarrow {}^4T_1(^4P)$			4.64

also assigned to Cr^{3+}, and their f values, though larger, are reasonable. The somewhat larger f values are thought to be caused by the presence of the Fe^{2+} in the tetrahedral sites enhancing the transition probabilities of the Cr^{3+} in the octahedral sites. The chemically determined concentration of Fe in R97 is $(3.6\pm0.3)\times10^{20}$ cm^{-3}. Thus, 1 out of every 42 tetrahedral sites is occupied by Fe. Since each tetrahedral site in the spinel lattice has 12 nearest octahedral sites, there will be 4.3×10^{21} cm^{-3} octahedral sites, or 1/7 of the total which have an Fe atom as a next-nearest neighbor. If the Cr atoms are randomly distributed over the octahedral sites, then 1/7 of them will be affected by the presence of the tetrahedral Fe in R97. For low concentrations of Fe and Cr the fraction of the Cr sites that are paired with Fe is proportional to the Fe concentration. The fraction of the Cr sites in R54 that are paired is 0.004. The optical absorption spectrum of the Cr^{3+} in R54 is thus not influenced by the presence of the Fe^{2+}. In R97 the case is different. The absorption of the 2.24 and "3.29" eV bands is apparently enhanced. Also the "3.29" eV band has been split into two components, as mentioned previously. In addition there are 3 normally very weak octahedral Cr^{3+} peaks which show up in Fig. 7 as absorption bands at 1.89, 2.72, and a shoulder at 3.7 eV. The 1.89- and 2.72-eV bands have been seen before in blue or blue-green, natural spinel crystals[64,68] which are known[66] to contain Fe^{2+}. The 3.7-eV band has been seen before by Weigel[63] in synthetic $MgAl_2O_4$ containing Cr^{3+}. The assignment of these several bands to various known transitions in octahedral Cr^{3+} is given in Table VIII. The corresponding energies found for $[Cr(H_2O)_6]^{3+}$ are also shown.[73] The agreement is good, indicating that the several absorption bands in R97 between 1.5 and 4.0 eV are caused by the Cr^{3+}. None of them can be attributed directly to optical transitions within the d shell of the tetrahedral Fe^{2+}. This is in contrast to the suggestions of Schlossmacher[64–66] who believed that the absorption spectrum in the visible range of the Fe containing $MgAl_2O_4$ crystals was actually due to the Fe^{2+}. In his crystals as in R97, it is caused by the presence of Cr^{3+}. The Fe^{2+} has, however, strongly modified the Cr^{3+} absorption spectrum. The absorption band at 1.34 eV in R97 is at too low an energy to be caused by the octahedral Cr^{3+}. As explained before, it is believed to be associated with the presence of some tetrahedral Fe^{3+}.

[73] Y. Tanabe and S. Sugano, J. Phys. Soc. Japan 9, 766 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Low-Temperature Galvanomagnetic Phenomena in an Intense Electric Field

HERBERT F. BUDD*
Laboratoire de Physique de l'Ecole Normale Superieure, Paris, France
(Received 2 December 1963; revised manuscript received 16 January 1964)

The electron distribution function is calculated for a many-valley semiconductor in an intense electric field for the case of anisotropic scattering by acoustical phonons at low temperatures.

I. INTRODUCTION

IN this paper we shall consider the low-temperature galvanomagnetic properties of a semiconductor in an intense electric field. In order to calculate the low-temperature anisotropic transport properties of semiconductors such as n-type germanium and silicon, it is necessary to abandon both the simple picture of spherical constant energy surfaces and the assumption of phonon equipartition.

Zylberstejn and Conwell[1] have studied the deviations of the phonon distribution from equilibrium in a semiconductor in an intense electric field at low temperatures. They show that the deviation is small for a sample of small dimensions and small carrier concentration. We shall treat this latter case and assume that the phonon distribution remains in equilibrium.

The predominant electron scattering mechanism in a pure semiconductor at low temperatures is by acoustical phonons and the equilibrium number of such phonons of crystal momentum $\mathbf{q}$ is given by

$$N_q = 1/(e^{qs/kT}-1), \quad (1)$$

where s is the velocity of sound, k is Boltzmann's constant, and T is the absolute temperature. The conservation of energy and crystal momentum require that the acoustical phonons with which an electron of energy ϵ can interact have a maximum energy of the order of $2(\epsilon ms^2)^{1/2}$ where m is the effective electron mass. At high temperatures the average electron energy $\bar{\epsilon}$, even in an intense electric field, is such that

$$2(\bar{\epsilon}ms^2)^{1/2} \ll kT \quad \text{and therefore} \quad qsN_q \approx kT. \quad (2)$$

This is the case of equipartition of energy.

At low temperatures, on the other hand, an intense electric field may result in the average electron energy $\bar{\epsilon}$ being sufficiently large such that

$$2(\bar{\epsilon}ms^2)^{1/2} > kT \quad \text{and} \quad N_q \sim \exp-\left(\frac{2(\bar{\epsilon}ms^2)^{1/2}}{kT}\right). \quad (3)$$

Since the probability of phonon emission and absorption are proportional to $1+N_q$ and N_q, respectively, we may in this case neglect N_q compared to one and consider only acoustical phonon emission; this is the case of zero-point scattering.

Stratton[2] has calculated the electron distribution function in the limit of zero-point scattering for the case of spherical constant energy surfaces and isotropic scattering, while Paranjape[3] has calculated the electron temperature for a Boltzmann distribution for the same case. Koenig, Brown, and Schillinger (K.B.S.)[4] have recently applied Shibuya's[5] theory of hot electrons in order to describe hot electron phenomena in n-type germanium at low temperatures. Shibuya's treatment is based on the assumption of phonon equipartition and is therefore not applicable under the condition for which K.B.S. have used it. We see in fact from Table II of their paper that the electron temperatures, or the corresponding mean electron energies, do not satisfy the equipartition condition (2).

Stratton's zero-point distribution function and Paranjape's calculation predict an electron mobility $\mu \sim E^{-0.8}$, where E is the electric field. This is in agreement with the low-temperature conductivity measurements of Bray and Brown.[6]

In this paper we shall extend the previous calculations[2,3] by considering a many-valley semiconductor with ellipsoidal constant energy surfaces in an intense electric field and a magnetic field. We treat the case of acoustical phonon scattering in the zero-point limit, and allow for anisotropic scattering in our calculation.

In Sec. II of this paper we set up the Boltzmann equation for a many-valley semiconductor, and in Sec. III we derive expressions for the energy and momentum relaxation in the case of anisotropic zero-point scattering. We solve the Boltzmann equation for the distribution function in Sec. IV.

II. BASIC EQUATIONS

In a previous publication,[7] the author has shown that the Boltzmann equation for the case of ellipsoidal constant energy surfaces can be expressed as two coupled equations for S and A, the isotropic and aniso-

* Present address: Xerox Corporation, Rochester, New York.
[1] A. Zylberstejn and E. Conwell, Phys. Rev. Letters 11, 417 (1963).
[2] R. Stratton, Proc. Roy. Soc. (London) 242, 355 (1957).
[3] B. V. Paranjape, Proc. Phys. Soc. (London) B70, 628 (1959).
[4] S. Koenig, R. Brown, and W. Schillinger, Phys. Rev. 128, 1668 (1962).
[5] M. Shibuya, Phys. Rev. 99, 1189 (1955).
[6] R. Bray, D. Brown, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechoslovakian Academy of Sciences, Prague, 1961).
[7] H. Budd, Phys. Rev. 131, 1520 (1963).

CC-0. In Public Domain. Gurukul K... Collection

Digitized by Arya Samaj Foundation Chennai and eGangotri

tropic parts of the distribution function, respectively,

$$e\mathbf{E}'\cdot\nabla_{p'}S+\mathrm{An}\{e\mathbf{E}'\cdot\nabla_{p'}A\}+\frac{e}{m_0}\mathbf{p}'\times\mathbf{B}'\cdot\nabla_{p'}A=\hat{C}'A\,,$$

$$\mathrm{Is}\{e\mathbf{E}'\cdot\nabla_{p'}A\}=\hat{C}'S\,, \tag{4}$$

with $\mathbf{p}'=\alpha\mathbf{p}$, $\mathbf{E}'=\alpha\mathbf{E}$, $\mathbf{B}'=R\mathbf{B}$, $\epsilon=p'^2/2m_0$,

$$\alpha=\begin{bmatrix}\frac{m_0}{m_x} & 0 & 0\\ 0 & \frac{m_0}{m_y} & 0\\ 0 & 0 & \frac{m_0}{m_z}\end{bmatrix}^{1/2} \qquad R=\begin{bmatrix}m_x & 0 & 0\\ 0 & m_y & 0\\ 0 & 0 & m_z\end{bmatrix}^{1/2}\frac{m_0}{(m_xm_ym_z)^{1/2}}\,, \tag{5}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors, ϵ and $\mathbf{p}$ are the electron energy and momentum and m_x, m_y, m_z are the electron masses corresponding to the three principal axes of the ellipsoid.

An$\{x\}$ and Is$\{x\}$ represent the anisotropic and isotropic parts of x respectively, and $\hat{C}'$ represents the collision operator in $\mathbf{p}'$ space.

Assuming that $\hat{C}'A$ can be described by an energy-dependent tensor whose principal axes coincide with those of the ellipsoid:

$$\tau=\begin{bmatrix}\tau_x(\epsilon) & 0 & 0\\ 0 & \tau_y(\epsilon) & 0\\ 0 & 0 & \tau_z(\epsilon)\end{bmatrix}, \tag{6}$$

and neglecting the An$\{e\mathbf{E}'\cdot\nabla\rho'A\}$ term (see Appendix II), we obtain straightforwardly:

$$A=-\frac{dS}{d\epsilon}\mathbf{V}'\cdot\left[\frac{\tau(e\mathbf{E}'+\tau e\mathbf{E}'\times e\mathbf{B}'/m_0)+(e^2/m_0^2)\mathbf{B}'(\mathbf{B}'\cdot e\mathbf{E}')\bar{\tau}^3}{1+(e^2/m_0^2)\bar{\tau}^3[(B_x'^2/\tau_x)+(B_y'^2/\tau_y)+(B_z'^2/\tau_z)]}\right], \tag{7}$$

$$\hat{C}'S=-\frac{2}{3m_0\epsilon^{1/2}}\frac{d}{d\epsilon}\left[\frac{dS}{d\epsilon}\epsilon^{3/2}\left(\frac{e^2\mathbf{E}'\tau\mathbf{E}'+\bar{\tau}^3(e^2\mathbf{E}'\cdot\mathbf{B}'/m_0)^2}{1+(\bar{\tau}^3e^2/m_0^2)[(B_x'^2/\tau_x)+(B_y'^2/\tau_y)+(B_z'^2/\tau_z)]}\right)\right], \tag{8}$$

where $\mathbf{V}'=(\mathbf{p}'/m_0)$ and $\bar{\tau}^3=\tau_x\tau_y\tau_z$.

We shall consider the solution of this equation after having calculated the energy relaxation term $\hat{C}'S$, and the relaxation time tensor τ.

III. CALCULATION OF ENERGY RELAXATION TERM AND RELAXATION TIME TENSOR

In this section we shall derive expressions for τ and $\hat{C}'S$ for a many-valley semiconductor with anisotropic scattering by the zero-point acoustical phonons. Denoting by M_L^2, c_L and M_T^2, c_T, the matrix elements and elastic constants corresponding to the longitudinal and transverse modes respectively, we obtain

$$\hat{C}f(\mathbf{p})=\frac{2\pi}{\hbar(2\pi\hbar)^3}\left\{\int d\mathbf{q}[f(\mathbf{p}+\mathbf{q})(1+N_q)-f(\mathbf{p})N_q][X_L+X_T]+\int d\mathbf{q}[f(\mathbf{p}-\mathbf{q})N_q-f(\mathbf{p})(1+N_q)][Y_L+Y_T]\right\}; \tag{9}$$

$$X_{L,T}=(qs_{L,T}/2c_{L,T})M_{L,T}^2\delta(\epsilon(\mathbf{p}+\mathbf{q})-\epsilon(\mathbf{p})-qs_{L,T})\,,$$

where

$$Y_{L,T}=(qs_{L,T}/2c_{L,T})M_{L,T}^2\delta(\epsilon(\mathbf{p}-\mathbf{q})-\epsilon(\mathbf{p})+qs_{L,T})\,,$$

$\rho s_L^2=c_L$, $\rho s_T^2=c_T$ and where ρ is the density. The matrix elements have been calculated by Herring and Vogt,[8] and we see from Table V of their paper that M_L^2 and M_T^2 are simple polynomials in $\cos^2\theta$, where θ is the angle between the principal axis of the ellipsoid and $\mathbf{q}$.

Transforming the elliposidal constant energy surfaces to spheres (5) in both the electron and phonon space,

[8] C. Herring and E. Vogt, Phys. Rev. **101**, 944 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

we obtain

$$\hat{C}'f=\frac{2\pi m_T m_L^{1/2}}{\hbar(2\pi\hbar)^3 m_0^{3/2}}\Big\{\int d\mathbf{q}'[f(\mathbf{p}'+\mathbf{q}')(1+N_{q^*})-f(\mathbf{p}')N_{q^*}][X_L'+X_T'] + \int d\mathbf{q}'[f(\mathbf{p}'-\mathbf{q}')N_{q^*}-f(\mathbf{p}')(1+N_{q^*})][Y_L'+Y_T']\Big\}, \quad (10)$$

where

$$X_{L,T}'=(q^*s_{L,T}/2c_{L,T})M_{L,T}^2(\cos^2\theta^*) \times\delta(\epsilon(\mathbf{p}'+\mathbf{q}')-\epsilon(\mathbf{p}')-q^*s_{L,T}),$$

$$Y_{L,T}'=(q^*s_{L,T}/2c_{L,T})M_{L,T}^2(\cos^2\theta^*) \delta(\epsilon(\mathbf{p}'-\mathbf{q}')-\epsilon(\mathbf{p}')+q^*s_{L,T}).$$

The principal axis of the ellipsoid has been taken as the polar axis in $\mathbf{q}'$ space; and m_T and m_L are the transverse and longitudinal masses, respectively;

$$q^*=(m_T/m_0)^{1/2}q'[1+(K-1)\cos^2\theta']^{1/2},$$

$$\cos\theta^*=\frac{K^{1/2}\cos\theta'}{[1+(K-1)\cos^2\theta']^{1/2}}, \quad K=\frac{m_L}{m_T}, \quad (11)$$

where θ' is the cone angle in $\mathbf{q}'$ space.

We first calculate $\hat{C}'S$, where S is a function of energy only. If the scattering were isotropic the matrix elements would be angularly independent and $\hat{C}'S$ would depend only on energy. In our case, M_L^2 and M_T^2 are angularly dependent and $\hat{C}'S$ will therefore in general be a function of the direction of $\mathbf{p}'$. We shall calculate the average of $\hat{C}'S$ over all directions of $\mathbf{p}'$, i.e.,

$$\langle\hat{C}'S\rangle_{\text{av}}=\int\int\hat{C}'Sd\Omega_{p'}\Big/\int\int d\Omega_{p'}, \quad (12)$$

where $d\Omega_{p'}$ is the differential solid angle in $\mathbf{p}'$ space.

Assuming that:

(1) The distribution function varies slowly over an energy interval equal to the phonon energy:

$$S(\epsilon+qs)=S(\epsilon)+qs\frac{dS}{d\epsilon}+\frac{(qs)^2}{2}\frac{d^2S}{d\epsilon^2}. \quad (13a)$$

(2) The average electron energy $\bar{\epsilon}$ is sufficiently large such that N_q is negligible compared to one:

$$2(\bar{\epsilon}ms^2)^{1/2}>kT \quad (13b)$$

and carrying out the required integrations we obtain

$$\hat{C}'S=\frac{64\pi^3m_T^2m_L^{1/2}}{\sqrt{2}h^4\rho\epsilon^{1/2}}\Sigma^{2*}\frac{d}{d\epsilon}\left[\epsilon^2\left(S+\frac{4\sqrt{2}}{5}m_T^{1/2}\epsilon^{1/2}V^*\frac{dS}{d\epsilon}\right)\right], \quad (14)$$

where

$$\Sigma^{2*}=\int_0^1R^2\left[M_L^2\left(\frac{Kx^2}{R^2}\right)+M_T^2\left(\frac{Kx^2}{R_2}\right)\right]dx,$$

$$V^*=\int_0^1R^3\left[s_LM_L^2\left(\frac{Kx^2}{R^2}\right)+s_TM_T^2\left(\frac{Kx^2}{R^2}\right)\right]dx\Big/\Sigma^{2*},$$

and

$$R=[1+(K-1)X^2]^{1/2}.$$

A relaxation time does not generally exist when the matrix elements are angularly dependent; we follow the procedure of Herring and Vogt and define two principal relaxation times as follows:

$$1/\tau_{11}=-\int\int\phi_{11}'\hat{C}'\phi_{11}'d\Omega_{p'}\Big/\int\int\phi_{11}'^2d\Omega_{p'},$$

$$1/\tau_{\perp}=-\int\int\phi_{\perp}'\hat{C}'\phi_{\perp}'d\Omega_{p'}\Big/\int\int\phi_{\perp}'^2d\Omega_{p'}, \quad (15)$$

with

$$\phi_{11}'=\mathbf{p}'\cdot\mathbf{a}/p', \quad \phi_{\perp}'=\mathbf{p}'\cdot\mathbf{b}/p',$$

where $\mathbf{a}$ and $\mathbf{b}$ are unit vectors parallel to the major and minor axes of the ellipsoid, respectively. We obtain the following expressions for the two relaxation times:

$$\frac{1}{\tau_{11,\perp}}=\frac{48(2\pi)^3}{5h^4}m_T^{3/2}m_L^{1/2}\epsilon N_{11,\perp}, \quad (16)$$

where

$$N_{11}=T_1, \quad N_{\perp}=\frac{T_0-T_1}{2},$$

and

$$T_n=\int_0^1dxR(x)x^{2n}\left[\frac{s_L}{c_L}M_L^2\left(\frac{Kx^2}{R^2}\right)+\frac{s_T}{c_T}M_T^2\left(\frac{Kx^2}{R^2}\right)\right].$$

The matrix elements and average elastic constants are given in Table V of Herring and Vogt's paper:

$$M_L^2\left(\frac{Kx^2}{R^2}\right)=\Sigma_d^2\left(1+\frac{Kx^2W}{R^2}\right)^2\left[1+\gamma_L\left(0.15-\frac{1.5Kx^2}{R^2}+\frac{1.75K^2x^4}{R^4}\right)\right],$$

CC-0. In Public Domain. Gurukul K

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$M_T{}^2\left(\frac{Kx^2}{R^2}\right)=\frac{Kx^2\Sigma_d{}^2W^2}{R^2}\left(1-\frac{Kx^2}{R^2}\right)\left(1+\gamma_T\frac{Kx^2}{R^2}\right);\quad W=\frac{\Sigma_\mu}{\Sigma_d},$$

$$c_L=c_{12}+2c_{44}+0.6c^*;\qquad \frac{1}{c_T}=\begin{cases}\dfrac{0.375}{c_{44}}+\dfrac{0.625}{c_{44}+\frac{1}{3}c^*}, & \text{(111) valley}\\[2ex] \dfrac{3}{c_{44}+\frac{1}{3}c^*}-\dfrac{2}{c_{44}+\frac{1}{2}c^*}, & \text{(100) valley,}\end{cases}$$

$$\gamma_L=\begin{cases}\dfrac{2c^*}{3c_L}, & \text{(111) valley}\\[2ex] -\dfrac{c^*}{c_L}, & \text{(100) valley}\end{cases}\qquad \gamma_T=\begin{cases}\dfrac{9}{8}c_T\left(\dfrac{1}{c_{44}}-\dfrac{1}{c_{44}+\frac{1}{3}c^*}\right), & \text{(111) valley}\\[2ex] 6c_T\left(\dfrac{1}{c_{44}+\frac{1}{2}c^*}-\dfrac{1}{c_{44}+\frac{1}{3}c^*}\right), & \text{(100) valley,}\end{cases}$$

where Σ_d and Σ_μ are the deformation potential constants for dilation and uniaxial shear. Inserting these matrix elements in Eqs. (16) and (14) we obtain the relaxation time tensor and $\hat{C}'S$ (see Appendix I).

IV. THE DISTRIBUTION FUNCTION

We shall now solve Eq. (8) for S, the isotopic part of the distribution function. Substituting Eq. (14) in Eq. (8), we obtain

$$S=\exp-\left[\int_0^\epsilon \frac{H\epsilon^{1/2}d\epsilon}{\dfrac{2}{3m_0}\left(\dfrac{e^2\mathbf{E}'\tau\mathbf{E}'+[(e^2/m_0)\mathbf{E}'\cdot\mathbf{B}']^2\bar{\tau}^3}{1+(e^2\bar{\tau}^3/m_0{}^2)[(B_x'{}^2/\tau_x)+(B_y'{}^2/\tau_y)+(B_z'{}^2/\tau_z)]}\right)+\bar{H}}\right],\tag{17}$$

where

$$H=\frac{64\pi^3m_T{}^2m_L{}^{1/2}\Sigma^{2*}}{\sqrt{2}h^4\rho};\quad \bar{H}=\frac{4\sqrt{2}}{5}m_T{}^{1/2}V^*\epsilon H.$$

We shall consider the case of a small magnetic field ($\omega^2\tau^2\ll1$) and we neglect the field-independent term in the denominator of the integrand of Eq. (17) (see Appendix II). We then obtain

$$S=\exp-(\epsilon/kT_e)^{5/2},\tag{18}$$

$$A=-\frac{dS}{d\epsilon}\mathbf{V}'\cdot\left[\tau\left(e\mathbf{E}'+e\tau\mathbf{E}'\times\frac{e\mathbf{B}'}{m_0}\right)\right],\tag{19}$$

with

$$T_e=X(\gamma)G\frac{[E\mu_aT^{3/2}/s_L]^{4/5}}{[m_Ts_L{}^2/k]^{1/5}}\left[\frac{3\bar{\Sigma}^4}{16\Sigma^{2*}N_1s_L\rho}\right]^{2/5},\tag{20}$$

where

$$X(\gamma)=\left[1+\gamma^2\left(\frac{1}{K'}-1\right)\right]^{2/5},\quad K'=K\frac{\tau_\perp}{\tau_{11}},$$

$$G=\left[\frac{\sqrt{2}\pi(27)(25)}{16^2(2+1/K)^2}\right]^{2/5},$$

and γ is the direction cosine of $\mathbf{E}$ with respect to the longitudinal axis of the ellipsoid. μ_a is the low-field mobility for acoustical phonon scattering and is related to the average deformation potential which appears in Herring and Vogt's paper [Eqs. (49) and (50)]

$$\bar{\Sigma}^4=[\xi\Sigma_d{}^2+\eta\Sigma_\mu\Sigma_d+\zeta\Sigma_\mu{}^2]^2.$$

When we consider the simple model of spherical constant energy surfaces, $\Sigma_\mu=0$ and neglect elastic anisotropy ($c^*=0$), the last factor in our expression for T_e becomes equal to 1, as does $X(\gamma)$, and Eq. (18) becomes identical with the distribution function calculated by Stratton.[2] Since the low-field mobility for acoustical phonon scattering varies as $T^{-3/2}$, our expression T_e is independent of temperature.

Using Eqs. (19) and (20) we obtain the following expressions for the mobility μ and the low magnetic-field Hall coefficient R.

$$\mu=PE^{-0.8};\quad R=B,$$

where P and B are complicated functions of the orientation of the electric and magnetic fields with respect to the crystallographic axes but are independent of the *magnitudes* of the applied fields. The field dependence of the mobility is in agreement with the experimental results of Bray and Brown.[6] A discussion of the orientation dependence will be considered in a future publication.

ACKNOWLEDGMENTS

The author expresses his sincere thanks to Professor P. Aigrain and Dr. E. Conwell for their stimulating discussions. Thanks are also due to Dr. A. Zylberstejn and D. Olechna for helpful suggestions and to the Xerox Corporation for partial support of this work.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX I

We shall express all the required integrals of Sec. III in terms of the elementary integrals

$$\langle m,n\rangle \equiv \int_0^1 \frac{x^m dx}{[1+(K-1)x^2]^{n/2}}, \tag{A1}$$

$$\Sigma^{2*}=\Sigma_d{}^2\{\tfrac{2}{3}+\tfrac{1}{3}K(1+W)^2+\gamma_L[0.1-0.45K+1.75K^2\langle 4,2\rangle+2KW(0.05-1.5K\langle 4,2\rangle+1.75K^2\langle 6,4\rangle)$$
$$+(KW)^2(0.15\langle 4,2\rangle-1.5K\langle 6,4\rangle+1.75K^2\langle 8,6\rangle)]+\gamma_T(KW)^2(\langle 4,2\rangle-K\langle 6,4\rangle)\}, \tag{A2}$$

$$V^*=\frac{\Sigma_d{}^2}{\Sigma^{2*}}\{s_L(\langle 0,-3\rangle+2KW\langle 2,-1\rangle+K^2W^2\langle 4,1\rangle)+s_TKW^2(\langle 2,-1\rangle-K\langle 4,1\rangle)$$
$$+\gamma_Ls_L[0.15\langle 0,-3\rangle-1.5K\langle 2,-1\rangle+1.75K^2\langle 4,1\rangle+2KW(0.15\langle 2,-1\rangle-1.5K\langle 4,1\rangle+1.75K^2\langle 6,3\rangle)$$
$$+(KW)^2(0.15\langle 4,1\rangle-1.5K\langle 6,3\rangle+1.75K^2\langle 8,5\rangle)]+\gamma_Ts_T(KW)^2(\langle 4,1\rangle-K\langle 6,3\rangle)\}, \tag{A3}$$

$$T_{\substack{0\\1}}=\Sigma_d{}^2\Big\{\frac{s_L}{c_L}(\langle {}^0_2,-1\rangle+2KW\langle {}^2_4,1\rangle+(KW)^2\langle {}^4_6,3\rangle)+\frac{s_T}{c_T}KW^2(\langle {}^2_4,1\rangle-K\langle {}^4_6,3\rangle)$$
$$+\gamma_L\frac{s_L}{c_L}[0.15\langle {}^0_2,-1\rangle-1.5K\langle {}^2_4,1\rangle+1.75K^2\langle {}^4_6,3\rangle+2KW(0.15\langle {}^2_4,1\rangle-1.5K\langle {}^4_6,3\rangle+1.75K^2\langle {}^6_8,5\rangle)$$
$$+(KW)^2(0.15\langle {}^4_6,3\rangle-1.5K\langle {}^6_8,5\rangle+1.75K^2\langle {}^8_{10},7\rangle)]+\gamma_T\frac{s_T}{c_T}(KW)^2(\langle {}^4_6,3\rangle-K\langle {}^6_8,5\rangle)\Big\}. \tag{A4}$$

The integrals $\langle m,n\rangle$ are all straightforward. For $K=1$, $\langle m,n\rangle=1/(m+1)$, and for $K\neq 1$ one can easily show that

$$\langle m+2,n+2\rangle=\frac{1}{n(K-1)}\left[(m+1)\langle m,n\rangle-\frac{1}{K^{n/2}}\right] \tag{A5}$$

for $n\neq 0$. We have tabulated some of the integrals in Table I. The remaining ones are easily calculated by means of Eq. (A5).

TABLE I. Integrals.

	$K>1$	$K=1$	$K<1$
$\langle 0,-1\rangle$	$\frac{K^{1/2}}{2}+\frac{\sinh^{-1}(K-1)^{1/2}}{2(K-1)^{1/2}}$	1	$\frac{K^{1/2}}{2}+\frac{\sin^{-1}(1-K)^{1/2}}{2(1-K)^{1/2}}$
$\langle 2,-1\rangle$	$\frac{1}{8(K-1)}\left[K^{1/2}(2K-1)-\frac{\sinh^{-1}(K-1)^{1/2}}{(K-1)^{1/2}}\right]$	$\frac{1}{3}$	$\frac{1}{8(1-K)}\left[-K^{1/2}(2K-1)+\frac{\sin^{-1}(1-K)^{1/2}}{(1-K)^{1/2}}\right]$
$\langle 4,2\rangle$	$\frac{1}{(K-1)^2}\left[\frac{K-4}{3}+\frac{\tan^{-1}(K-1)^{1/2}}{(K-1)^{1/2}}\right]$	$\frac{1}{5}$	$\frac{1}{(1-K)^2}\left[\frac{K-4}{3}+\frac{\tanh^{-1}(1-K)^{1/2}}{(1-K)^{1/2}}\right]$

APPENDIX II

We see from Eq. (18) that $\bar{\epsilon}=kT_e/\Gamma(0.6)\approx\frac{3}{2}kT_e$ and therefore using Eq. (3) the zero-point scattering condition is satisfied for electric fields large enough such that

$$\sqrt{6}\left(\frac{m^*s^{*2}}{k}\right)^{1/2}(X(\gamma)G)^{1/2}\frac{[E\mu_aT^{3/2}/s_L]^{2/5}}{[m_Ts_L{}^2/k]^{1/10}}\left[\frac{3\Sigma^4}{16\Sigma^{2*}N_1s_L\rho}\right]^{1/5}>T, \tag{A6}$$

where m^* is a complicated average of the longitudinal and transverse masses and is given approximately by

$$m^*\approx\tfrac{2}{3}m_T(K^{3/2}-1)/(K-1)$$

and s^* is the average sound velocity.

For spherical constant energy surfaces and $\Sigma_\mu=c^*=0$, Eq. (20) becomes approximately

$$2.4(ms^2/k)^{2/5}[E\mu_aT^{3/2}/s]^{2/5}>T.$$

The neglect of the field-independent term in the denominator of Eq. (17) is equivalent to neglecting ms^2/kT_e compared to one. It is easily seen that the $An\{e\mathbf{E}'\cdot\nabla_{p'}A\}$ term is a correction to the distribution function due to neglecting higher order terms than the first-order spherical harmonics, i.e., terms of the form $Y_{2n}(\theta,\phi)$, etc.

CC-0. In Public Domain. Gurukul K...

Digitized by Arya Samaj Foundation Chennai and eGangotri

Spin-Lattice Relaxation Time of Conduction Electrons in Sodium Metal*

FREDERICK VESCIAL,† N. S. VANDERVEN, AND R. T. SCHUMACHER‡
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received 6 January 1964)

We have measured the temperature dependence of the spin-lattice relaxation time T_1 in sodium metal from 14 to 77°K. Our results agree with those of Feher and Kip, and Levy, in the range 20 to 40°K, but disagree with Feher and Kip above 40°K. As a consequence, we find marked deviation from the relation T_1T=constant even at the highest temperatures of our measurements. The temperature dependence we observe lends strong support to the recent calculation of Yafet in the temperature range 35 to 77°K.

I. INTRODUCTION

THE question of the temperature dependence of the conduction-electron relaxation time in metals has recently been re-examined by Yafet.[1] He finds that in the very low-temperature limit, $T \ll \theta_D$, where θ_D is an appropriate Debye temperature for the lattice, the spin-lattice relaxation time T_1 should vary as T^{-5}, contrary to the previous expectations that T_1T=constant over the entire high- and low-temperature range. The only metal in which a temperature-dependent T_1 has been observed over an appreciable range of temperatures is sodium. In that metal Feher and Kip[2] measured T_1 from 300 to 4°K and concluded that their data supported the relation $T_1T \cong 3.0\times10^{-6}$ sec-°K over the entire range of their measurements. In fact, their published data do not support this conclusion below 55°K, nor do the data of Levy.[3] All these investigators show longer T_1's in the range 20 to 40°K than predicted by $T_1T \cong 3.0\times10^{-6}$ sec-°K. Narrower lines than expected at 20°K were also observed in this laboratory during a recent remeasurement of the sodium conduction-electron susceptibility.[4]

We report here a remeasurement of $T_1(T)$ which supports the conclusion of Yafet[1] that $T_1(T)$ deviates from the $1/T$ dependence as T becomes less than an appropriate Debye temperature for the sodium lattice. We describe those aspects of the apparatus relevant to this particular measurement, some precautions taken in sample preparation and experimental measuring technique, and we compare our results to the conclusions of Yafet.

II. EXPERIMENTAL DETAILS

A. Microwave Apparatus

The measurements were made entirely at X band with a single klystron superheterodyne spectrometer described previously.[5] The microwave cavity was cylindrical, and was operated in the TE_{011} mode, which requires that the sample be inserted into the geometric center of the cavity. In this configuration the microwave field H_1 is vertical in the laboratory and the dc field, H_0, is in the horizontal plane. The cavity was matched by a variable coupler of the type described by Gordon.[6]

The samples, which will be described in detail below, were contained in the innermost of a series of concentric Pyrex tubes. The sample tube was 3 mm o.d., a few centimeters long and was completely sealed. The next tube, 5 mm o.d., was approximately one meter long. It was inserted into a tube 8 mm o.d., which is permanently attached to the microwave system. The 8 mm o.d. tube extends halfway into the cavity, is coaxial with it and sealed at the lower end, and in the other direction extends to the top of the Dewar system, where it forms part of an "air-lock" system similar to one described by Estle[7] Samples were changed while cryogenic liquid was still in the Dewar system by removing the 5 mm o.d. tube.

B. Cryogenic Apparatus and Temperature Measurements

The relative bulkiness of the cylindrical cavity necessitated the use of a Dewar system whose outer diameter (that of the nitrogen Dewar) is 5 in. The inner diameter of the inner Dewar is greater than $2\frac{1}{2}$ in. The cryogenic liquid used in our measurements was in all cases liquid hydrogen. However, the hydrogen is entirely excluded from the interior of the microwave cavity and the inside of the waveguide. The outermost pyrex tube is also not in direct contact with the hydrogen, but is separated from it by a cupronickel alloy sleeve which fits snugly around the Pyrex and is soldered to the top of the cavity. The sample is thus removed from direct contact with the bath by a metal sleeve and three Pyrex tubes, all of which nest together rather snugly. All the tubes except the sample tube were filled with a fraction of an atmosphere of He gas to improve the thermal contact between the sample and the bath.

* Supported by grants from the National Science Foundation and the Alfred P. Sloan Foundation. From a thesis submitted by Frederick Vescial in partial fulfillment of requirements for the Ph.D. degree.
† Present address: Autonetics, Inc., Anaheim, California.
‡ Alfred P. Sloan Foundation Fellow, 1959–1963.
[1] Y. Yafet, Solid State Phys. 14, 1 (1963).
[2] G. Feher and A. F. Kip, Phys. Rev. 98, 337 (1955).
[3] R. A. Levy, Phys. Rev. 102, 31 (1956).
[4] R. T. Schumacher and W. E. Vehse, Phys. Chem. Solids 24, 297 (1963).
[5] J. L. Hall and R. T. Schumacher, Phys. Rev. 127, 1892 (1962); also J. L. Hall, thesis, Carnegie Institute of Technology, 1961 (unpublished).
[6] J. P. Gordon, Rev. Sci. Instr. 32, 658 (1961).
[7] T. L. Estle and G. K. Walters, Rev. Sci. Instr. 32, 1058 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Temperature measurements were made with a nominally 390-Ω carbon resistance thermometer. This thermometer was calibrated from 4.2 to 77°K by a platinum resistance thermometer which had been calibrated by the National Bureau of Standards. The carbon thermometer became rather insensitive above 50°K, but was still adequate for our measurements. The thermometer could not be placed in direct contact with the sample, which was in any event sealed in the 3-mm-o.d. tube. The top of the sample tube projected just above the top of the cavity, and the carbon thermometer was pressed into direct physical contact with the tube at that point. We relied on the physical contact plus the helium exchange gas to keep the thermometer at the same temperature as the sample. In the region from 14 to 20°K there was always adequate agreement between the temperature of the bath as deduced by vapor pressure measurements and the temperature indicated by the thermometer. Since the thermometer was not in any poorer contact with the bath than was the sample, we have no reason to believe that the sample temperature differed in any significant way from that registered by the thermometer.

Above the boiling point of liquid hydrogen data were taken by allowing the system to warm gradually after the liquid hydrogen level dropped below the bottom of the microwave cavity. During this time the field was swept back and forth through the resonance while the derivative of the resonance line was continuously recorded as was the temperature of the carbon resistance thermometer. This procedure gave data at about every 2°K, except for a few intervals during which the microwave system was being retuned. The warmup procedure from 20 to 70°K took about 60 min. Our best estimate of the type of error introduced by this procedure is that the sample was warming faster than the thermometer, and that the error became worse as the temperature became higher. Errors in temperature measurement are probably responsible for the increased scatter in the data above 50°K.

C. Sample Preparation

Our sodium samples were prepared by the technique introduced by Levy.[3] Briefly, the technique is to prepare dilute sodium-ammonia solutions, and then freeze them. Metallic sodium precipitates out upon freezing in the form of dispersed small particles. We used sodium, kindly given to us by Dr. Raymond Bowers of Cornell University, which had a resistivity ratio of more than 7000. We found it advisable to dissolve this sodium in ammonia which had first been cleaned by dissolving some normal purity sodium in it. This procedure removes any impurities in the ammonia which might react with sodium. The samples were confined to a length in the 3-mm-o.d. tube of about 2 mm in order to confine the sample to the smallest feasible volume to minimize effects of magnetic field inhomogeneity. After the required amount of ammonia had been introduced, the sample was frozen in liquid nitrogen and sealed off under vacuum a few centimeters above the sample. The 3-mm-o.d. tube was inserted in a long 5-mm-o.d. tube and stored under liquid nitrogen.

D. Some Experimental Precautions

The relaxation times which we measured were in the range from 5×10^{-8} to 5×10^{-7} sec. Such relaxation times are awkward to measure directly, so we took advantage of the equality of the transverse and longitudinal relaxation times[8] and measured the linewidth. Such a procedure is line shape-dependent, so it will be given extended attention below. Here we want to point out that at low temperatures, even for relatively short relaxation times, a serious danger to accurate measurement lies in partial saturation of the resonance line. The danger arises not from ordinary saturation broadening, but from the Overhauser effect.[8,9] We have observed at temperatures as high as 20°K effects predicted by Kaplan[10] and observed by Gueron and Ryter.[11] If the external field is swept slowly from above resonance while the rf field is strong enough to saturate the resonance even partially, the sodium nuclei will become partially polarized through the Overhauser effect. They will in turn contribute to the total field seen by the electrons an amount H^* which depends on the microwave power and the size of the sodium particles. Thus, the electrons still satisfy the resonance condition with the applied field an amount H^* below the resonance field, and absorption is still observed.

If large, the effect is readily identifiable, and it is simply required, in order to observe the natural linewidth, that the microwave power be sufficiently small. Our klystron operated at approximately 50 mW. We were able to observe Overhauser polarization at 20°K and below in this manner with powers incident on the cavity which were 20 dB down from the klystron power. Our linewidth measurements were made with powers 40 dB down from 50 mW, and we required power-independent linewidths over a 10-dB range of powers before the data was considered acceptable. Such low power levels required the full sensitivity of the superheterodyne system. This system[5] stabilizes the klystron to the sample cavity, so the single klystron superheterodyne system cannot be instantaneously retuned as the cavity and sample warm continuously above 20°K towards 77°K. Fortunately the relaxation time is short enough above 30°K to eliminate the saturation problem, and a simpler homodyne detection at higher power could be used. In this system the signal at the modulation frequency is taken directly from the

[8] T. R. Carver and C. P. Slichter, Phys. Rev. **102**, 975 (1956).
[9] A. W. Overhauser, Phys. Rev. **89**, 689 (1953).
[10] J. I. Kaplan, Phys. Rev. **99**, 1322 (1955).
[11] M. Gueron and C. Ryter, Phys. Rev. Letters **3**, 338 (1959).

CC-0. In Public Domain. Gurukul K...

Digitized by Arya Samaj Foundation Chennai and eGangotri

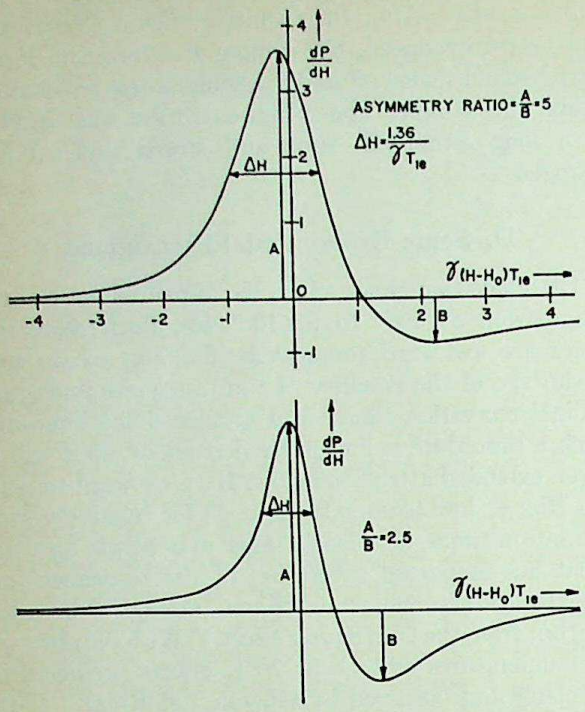


FIG. 1. (a) Derivative of ideal Dyson absorption line in the region of the anomalous skin effect. (b) Typical absorption derivative observed in our samples. Horizontal scale is not the same as in (a).

microwave crystal detector. Biasing power is provided by a small amount of microwave power reflected from the cavity, which is deliberately mismatched. The klystron is still locked to the microwave cavity, and retuning the system as the cavity frequency drifts takes only a few seconds.

We might remark parenthetically that we have made intensive measurements of the Overhauser shift in the hydrogen range, where the shift H^* becomes as large as 5 G. The sodium heat capacity in this region is large enough to prevent the metal particle temperature from rising above the bath temperature because of the microwave power absorbed by the spin system. Difficulties of this type have plagued Overhauser shift measurements on sodium in the helium region for years.[12] Unfortunately, direct measurements of the shift sufficiently precise to be of interest in themselves proved impossible, and our data was only good enough to be in qualitative agreement with the magnitude of H^* as expected from spin susceptibility[4] and Knight shift[13] measurements.

III. RESULTS AND DISCUSSION

In most of the temperature region over which our measurements were made, the conductivity of the sample is in the anomalous region; i.e., the electronic mean free path Λ is larger than the classical skin depth δ. In addition most of the sodium particles were large compared to Λ. The resonance absorption line shape for this case has been calculated by Kittel[2] based on the calculation by Dyson.[14] The derivative of the power absorbed versus $\gamma(H-H_0)T_1$ in Kittel's theory is plotted in Fig. 1. The line shape is characterized by extreme asymmetry, the ratio of the low-field peak to the high-field peak being given by $A/B=5$. The full width at half-maximum of the low-field peak is given by

$$\Delta H = 1.36/\gamma_e T_1, \tag{1}$$

where γ_e is the electron gyromagnetic ratio.

The line shape given by our samples is shown in the lower half of Fig. 1. The asymmetry is markedly less than 5:1, and was in fact always about 2.5:1. We believe our lines did not have the Dyson-Kittel shape because our samples in fact contained a mixture of particle sizes, with a substantial contribution to the signal coming from particles whose dimensions are comparable to or smaller than Λ. For these particles the line shape would be Lorentz, with $A:B=1$. This rationalization of our line shape is supported by the large amount of data taken on the Overhauser shift in the hydrogen temperature region, where we observed in detail the diminution of the intensity of the shifted line as the shift increased. This effect and its relation to particle size is discussed by Ryter.[12]

For a symmetrical Lorentz line, the relation between T_1 and ΔH is given by

$$\Delta H = 1.22/\gamma_e T_1. \tag{2}$$

For simplicity we have used Eq. (1) in converting from ΔH to T_1. A comparison of the observed and theoretical line shapes, along with the relations (1) and (2), leads us to believe that use of (1) does not result in T_1's more than 5% high. Since the temperature dependence of T_1 rather than its absolute value is our main interest, the uncertainty is of no importance.

The results of our measurements on all samples that showed a temperature-dependent ΔH to 14°K are

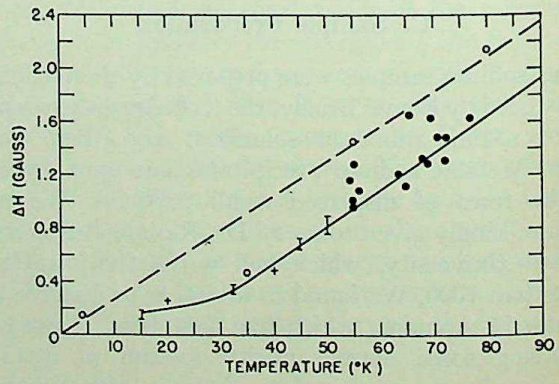


FIG. 2. ΔH versus T. ○, from Kip and Feher (Ref. 2); +, from Levy; ●, our data. Solid line and error flags explained in text. Dashed line is from the relation $T_1T=3\times10^{-6}$ sec −°K, converted to ΔH by Eq. (1) of text.

[12] Private communications from many investigators. See also C. Ryter, Phys. Rev. Letters 5, 10 (1960).

[13] B. R. McGarvey and H. S. Gutowsky, J. Chem. Phys. 21, 2114 (1953).

[14] F. J. Dyson, Phys. Rev. 98, 349 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

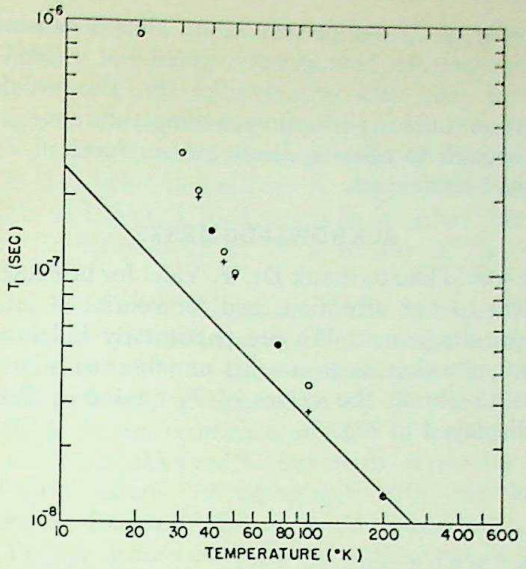


FIG. 3. The theoretical and experimental T_1's of Table I on a log-log plot. ○, Yafet; +, ●, experiment as in Fig. 2. The solid line is T_1T=constant drawn through theory and experiment at 200°K. The constant, 2.6×10^{-6} sec −°K, is not the same as in Fig. 2.

shown in Fig. 2, along with the data of Feher and Kip[2] and of Levy.[3] Approximately 25 data points below 20°K and 25 points between 30 and 50°K are represented by the smooth curve drawn through them and error flags giving the limits within which the data fall at several temperatures. The somewhat more scattered data above 50°K (twenty points) are shown as heavy dots. The straight line above 30°K is a least-squares fit to the data between 35 and 50°K; the equation of this line is

$$\Delta H=0.0277T-(0.577\pm0.010). \quad (3)$$

The data above 50°K is omitted from this fit not because of its scatter but because of line shape changes associated with the change between classical and anomalous skin effect regions.

The data, converted to T_1, is replotted in log-log form, in Fig. 3. The solid line is one drawn by Feher and Kip,[2] the dots, crosses, and circles are experiment and theory as explained below. We note several features of the results from Fig. 3. (1) Our data extrapolates at high temperatures towards the $T_1T=3.0\times10^{-6}$ sec-°K relation found by Feher and Kip. (2) If the theory of Yafet is fit to our data at 40°K it deviates significantly from T_1T=constant as high as 150°K. (3) Our data fit Yafet's theory from 77°K, our highest temperature, to 30°K, where our T_1's begin to fall well below the theory; and Yafet's theory, fit to our data at 40°K, agrees with Kip and Feher's data at high temperatures.

Feature (3) deserves discussion. The spin-lattice relaxation mechanism discussed by Yafet is an electron-phonon scattering accompanied by a spin flip. The mechanism which couples the two spin states is the spin-orbit interaction. Yafet's theory[1] has three parameters. The Debye phonon spectrum is assumed, but two Debye θ's are used, θ_N and θ_U. θ_N is the Debye temperature for longitudinal phonons; the N stands for "normal," since electron–longitudinal-phonon scattering is the normal resistivity scattering mechanism. umklapp processes are important because there is a term in the spin-flip matrix element involving only transverse phonons which vanishes for small electron scattering angle, but which contributes if the electron wave vector changes by the phonon wave vector plus a reciprocal lattice vector. Use of a different Debye θ for processes to which transverse phonons contribute corresponds to a modification of the strict Debye model which is realistic if θ_U is chosen less than θ_N. At least one transverse mode corresponding to a low θ has been observed in neutron diffraction by Woods *et al.*[15]

Dr. Yafet was kind enough to furnish us the details of his calculation, the results of which for $\theta_N=250$°K, $\theta_U=200$°K, are shown in Fig. 9 of his paper.[1] We were able to use his work to calculate the points shown in Fig. 3. The third adjustable parameter, the strength of the spin-flip matrix element, was chosen to fit the data at 40°K, and several values of θ_N and θ_U were tried. The results are quite sensitive to θ_U, with values of the pair (θ_N, θ_U) of (250°K, 80°K) and (250°K, 200°K), respectively, lying distinctly below and above the data. The points shown on Fig. 3 correspond to $\theta_N=250$°K and $\theta_U=150$°K. For $\theta_U=150$°K, longitudinal phonons contribute only from 7% to 15% to the total relaxation rate. It is worth pointing out that the Debye θ deduced from heat-capacity measurements is (156±2)°K.[16] Table I presents the calculated and measured T_1's at several temperatures.

Finally, we note that below 30°K the experimental T_1's deviate from theory; the data give $(T_1)_{\text{meas}} \cong \frac{3}{5}(T_1)_{\text{Yafet}}$ at 25°K, and the theoretical T_1 is beginning to rise very sharply just as the data are beginning to approach a temperature-independent value. Our line-

TABLE I. Calculated and measured relaxation times.

T in °K	Calculated[a] $T_1\times10^{-7}$ sec	Measured $T_1\times10^{-7}$ sec
20	8.85	...
30	3.00	...
35	2.13	2.00[b]
40	1.47	1.47[b]
45	1.22	1.12[b]
50	1.00	0.97[b]
75	0.52	0.52[b]
100	0.357	0.28[c]
150	0.223	0.22[c]
200	0.128	0.13[c]

[a] From Yafet (Ref. 1); fit to data at $T=40$°K, with parameters $\theta_U=150$°K, $\theta_N=250$°K.
[b] Calculated from Eq. (1), where ΔH was obtained from a least-squares fit of data between 35 and 50°K, as explained in text. Point at 75°K was also calculated from (1) and the least-squares ΔH.
[c] From Kip and Feher (Ref. 2), their Fig. 13.

[15] A. D. B. Woods, B. N. Brockhouse, R. H. March, and R. Bowers, Proc. Phys. Soc. (London) 79, 440 (1962).
[16] Douglas L. Martin, Phys. Rev. 124, 438 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

width in fact becomes temperature-independent at $\Delta H = 0.1$ G. A simple linewidth measurement of the proton resonance in a water sample of the same dimensions as the sodium sample showed the field inhomogeniety over the sample to be about 0.1 G. The field was so inhomogeneous over the small sample because the ratio of the magnet gap to the pole cap diameter was 5:12; the large gap was required in order to fit in the rather bulky microwave cavity and the glass Dewars. Thus our data may not be compared to Yafet's calculations below about 35°K, where the measured linewidths is 0.4 G. We can say in defense of our poor homogeneity that even if Yafet's calculations are correct only down to 20°K the linewidth would be on the order of $\frac{1}{10}$ our present limit. Thus a substantial improvement in homogeneity would not extend our range of valid data very significantly. The calculated relaxation times in the hydrogen temperature region are long enough to allow a direct measurement of T_1 by transient techniques.

ACKNOWLEDGMENTS

We would like to thank Dr. Y. Yafet for bringing this problem to our attention, and for continued interest and encouragement. We are particularly indebted to him for allowing us to use his unpublished numerical results to obtain the values of T_1 quoted in Table I and displayed in Fig. 3.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Absence of Pronounced Quadrupole Effects in the Nuclear Resonance of In^{115} in a Noncubic Environment*

LAWRENCE H. BENNETT AND REX J. SNODGRASS†
Metallurgy Division, National Bureau of Standards, Washington, D. C.
and
Department of Physics and Astronomy, University of Maryland, College Park, Maryland

(Received 23 December 1963)

The nuclear magnetic resonance of In^{115} has been observed throughout the cubic phase of Pb-In alloys, at 77 and 300°K. The Knight shift is 0.86% and is essentially unchanged with composition. The linewidths increase rapidly with In concentration, probably mostly due to pseudodipolar interaction. The average electric field gradient at a In^{115} nucleus for any configuration of Pb neighbors is about $q \sim 3 \times 10^{22}$ cm^{-3}, about 20 times smaller than for Cu having a Zn second neighbor in the fcc Cu lattice. The small value of the field gradient is discussed in terms of the behavior of the charge oscillations, and is related to the small change in Knight shift for Pb^{207} in these same alloys.

I. INTRODUCTION

THE nuclear electric-quadrupole moment interacts with the electric field gradient at the site of the nucleus. In a position of cubic symmetry, the electric field gradient vanishes and there is no quadrupole interaction. In the liquid phase of a metal, atomic motion is usually sufficiently rapid to average out the field gradients. Knight[1] pointed out that, in some cases, strong quadrupole interactions may obliterate the nuclear magnetic resonance (NMR) signal of a quadrupolar nucleus in the noncubic lattice of the pure metal, but that the nucleus may still exhibit a strong NMR signal when substituted in the cubic environment of another metal. For example, in pure orthorhombic Ga, the nuclear magnetic resonance has been observed in the liquid phase only. However, when Ga is substituted in the face-centered cubic lattice of Cu, the cubic environment substantially reduces the quadrupole coupling for many of the Ga atoms and the nuclear magnetic resonance is observable. Knight suggested that it might be possible to make suitable alloys of other metals for which the resonances are unobservable in the pure state.

Following this suggestion, one of the present authors[2] observed a strong NMR signal of In^{115} substituted in the cubic lattice of a 95% Pb−5% In alloy. If the indium is randomly distributed throughout the lattice, about one-half of the indium atoms have only lead for first, second, and third nearest neighbors. The intensity of the In^{115} resonance compared to the Pb^{207} in this alloy led to the conclusion that *these* atoms were observed, together with some of the satellite intensity. An alternative possibility is that all In^{115} nuclei participate in the resonance but that some of the satellite intensity is smeared out.

In order to distinguish between these possibilities the

* Some of this work forms a part of a Ph.D. thesis submitted to the University of Maryland by R. J. S. The experiments were performed at the National Bureau of Standards.

† Guest worker (1963–64), University of Paris, Faculté des Sciences, Orsay (Group II), Service de Physique des Solides, Orsay (Seine et Oise), France.

[1] W. D. Knight, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1956), Vol. 2. For further information on nuclear quadrupole effects in solids, see M. H. Cohen and F. Reif, *ibid.*, Vol. 5 (1957).

[2] L. H. Bennett, Bull. Am. Phys. Soc. 4, 251 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In^{115} resonance was looked for in other alloying proportions. Very surprisingly, the In^{115} resonance remained observable with almost no change in intensity per atom throughout the cubic lead phase.[3] Figure 1, for example, shows the observed room-temperature Pb^{207} and In^{115} NMR absorption derivatives in alloys containing 0.5 at.% In and 35 at.% In. There is no essential difference in the In^{115} resonance at 77°K in any of the alloys studied.

Since the start of this research, the pure quadrupole resonance of In^{115} in tetragonal indium metal has been discovered[4,5] with a room-temperature quadrupole coupling constant of approximately 30 Mc/sec. The high-field nuclear resonance of In^{115} in indium metal has also been observed,[6,7] but with extremely large quadrupole effects in agreement with the known quadrupole coupling constant in the tetragonal lattice. The In^{115} resonance has also been observed in liquid indium,[8] but here the rapid motion washes out quadrupole interactions. Nuclear magnetic resonance studies[9] of In^{115} in an unstrained single crystal of InSb revealed negligible quadrupole interactions, due to the high symmetry of the zinc blende lattice of InSb. Thus, none of these other results is directly comparable with our surprising observation of the absence of pronounced quadrupole effects in the noncubic environment of the nondilute Pb-In alloy.

In Sec. II we present the experimental data on the In^{115} NMR absorption derivatives linewidths, centers and relative intensities as a function of concentration of In in Pb, magnetic field, and temperature. In Sec. III, we discuss the significance of the observation of the In^{115} resonance in the "noncubic" environment of the high indium concentration alloys in lead. We relate this absence of pronounced quadrupole effects to the small changes in Knight shift in the Pb^{207} resonance in these same samples discussed in a preceding paper.[10]

II. EXPERIMENTAL METHODS AND RESULTS

Indium, a column III element, is dissolved in lead, a column IV element. Lead is a face-centered-cubic metal with melting point of 327°C. All measurements were made in the cubic phase which persists up to about 68 at.% In. In^{115} is 96% abundant, has a nuclear spin $I=9/2$, a large nuclear magnetic dipole moment (5.5 n.m.), and a large electric quadrupole moment (1.16 b).

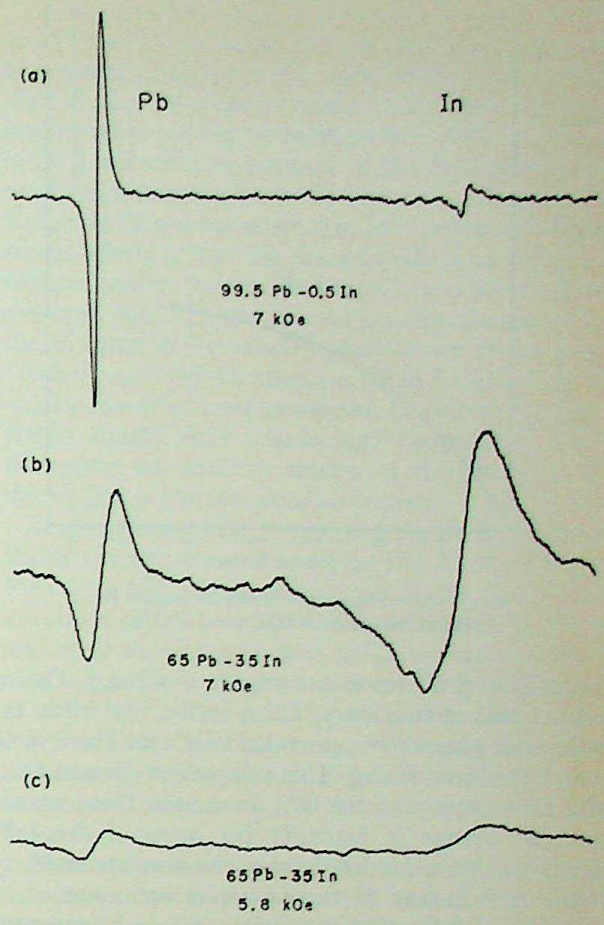


FIG. 1. First derivative of the room-temperature nuclear magnetic resonance absorptions of Pb^{207} and In^{115} in (a) 99.5% Pb−0.5% In at 7.0 kOe, (b) 65% Pb−35% In at 7.0 kOe and (c) 65% Pb−35% In at 5.8 kOe.

All of the resonances were observed at both 77° and 300°K with a Varian wide-line spectrometer. The experimental techniques and the preparation of the alloys have been discussed previously.[10,11] Typical resonances are shown in Fig. 1. Since the Pb^{207} and In^{115} resonance differ in shape and width, it is somewhat misleading to compare the recorded absorption derivatives in order to determine relative intensity. Instead, a point by point integration was made of each resonance, and the relative areas of the resulting absorption curves were compared. The assumption was made that the absorption is zero in the wings when the derivative is zero. This assumption has the effect of ignoring flat wings, and thus underestimating the contribution of In^{115} satellites.

If we make the reasonable assumption that we are observing all of the Pb^{207} nuclei, then a comparison of the relative intensities of the Pb^{207} and In^{115} resonances *in the same sample* gives an indication of the portion of

[3] L. H. Bennett and R. J. Snodgrass, Bull. Am. Phys. Soc. 7, 228 (1962).
[4] R. R. Hewitt and W. D. Knight, Phys. Rev. Letters 3, 18 (1959).
[5] W. W. Simmons and C. P. Slichter, Phys. Rev. 121, 1580 (1961).
[6] A. H. Silver (private communication).
[7] D. R. Torgeson and R. G. Barnes, Phys. Rev. Letters 9, 255 (1962).
[8] C. P. Flynn and E. F. W. Seymour, Proc. Phys. Soc. (London) 76, 301 (1960).
[9] R. G. Shulman, B. J. Wyluda, and P. W. Anderson, Phys. Rev. 107, 953 (1957).
[10] R. J. Snodgrass and L. H. Bennett, Phys. Rev. 132, 1465 (1963).
[11] R. J. Snodgrass and L. H. Bennett, Appl. Spectry. 17, 53 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

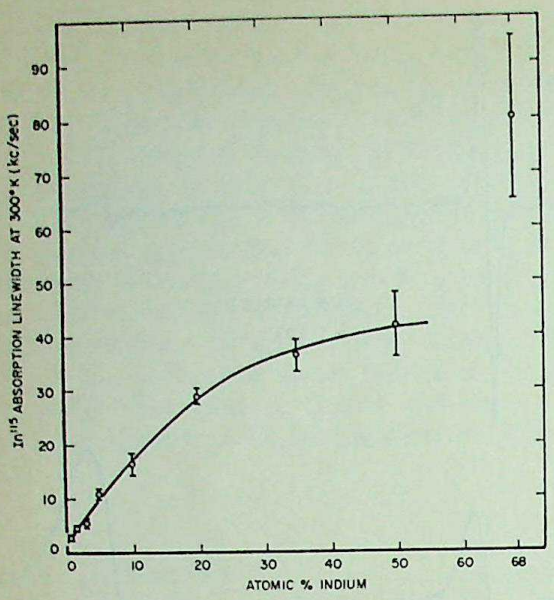


FIG. 2. In^{115} absorption linewidth versus In concentration at 300°K and 7.0 kOe.

the possible In^{115} resonance we are observing. This is independent of frequency, filling factor, and other experimental parameters, provided only that there is no modulation broadening. This comparison showed that, with the exception of the 68% In sample, there was almost no change in intensity per atom of the In^{115} throughout the cubic lead phase. The absolute intensity of the In^{115} in any of these samples corresponded to about 10 to 20% of that possible, in good agreement with the intensity expected from the central transition alone. We conclude that we are observing only the central line of essentially all In^{115} nuclei in the alloys, at both 77 and 300°K.

There is virtually no second-order quadrupole shift at 7.0 kOe or above. For the 35% In sample, Fig. 1[(b) and (c)] shows the room-temperature resonance absorption derivatives of In^{115} and Pb^{207} at 7.5 and at 5.8 kOe. The In^{115} resonance at the lower field shows some sign of second-order broadening. Any shift in center between these two In^{115} resonances (normalized to 5.8 kOe) is less than 1 kc/sec.

The presence of substantial ordering or clustering would of course change the main result of this paper, that we are observing only small quadrupole effects in a *noncubic* environment. An x-ray study[12] of the 20% In and the 35% In sample used in this research revealed no short-range order. The short-range order parameter in the ith shell around an A atom is defined as $\alpha_i=1-P_{AB_i}/X_B$, where P_{AB_i} is the probability of a B atom in the ith shell around an A atom, and X_B is the mole fraction of B. If α_1 was ≈-0.05 or greater in magnitude, it would have been detected. Clustering ($\alpha_1>0$) was not checked. A short-range order parameter equal to 0.14 was suggested recently.[13] Even if the short-range order were this high, fewer than 1% of the In would be in cubic environments in the 50-50 alloy.

The In^{115} linewidth as a function of In concentration is shown in Fig. 2. The increase in the In^{115} linewidth with In concentration can be understood if the pseudodipolar interaction is assumed to be about 10 times the ordinary dipolar. The pseudodipolar interaction was found to be about 10 times the dipolar for the Pb^{207} linewidth[10] in these same samples. No careful studies of the field dependence of the linewidth was made—it is believed that a combination of second-order quadrupolar broadening at the lower fields, and anisotropic and inhomogeneous isotropic Knight shift broadening[10] at the higher fields could be important. The 68% In point does not fall on the smooth curve in Fig. 2. From the Pb^{207} Knight shifts, we had previously[10] concluded that this alloy was in a two-phase region.

The resonance center is largely independent of concentration, occurring at 6.586±0.001 Mc/sec at 7000 Oe. This corresponds to a Knight shift of 0.86±0.01%, which is close to the value found in the tetragonal phase of pure In[7] (0.82±0.04% for the isotropic shift), or in the liquid phase of In metal[8] (0.79±0.03%, the ± taking into account the variation with temperature, and chemical shifts). There is no *a priori* reason for expecting this close agreement.

III. DISCUSSION

The most important result of this work is the observation of the In^{115} resonance in the "noncubic" environment of nondilute Pb–In alloys. This result was unexpected in view of the large quadrupole moment of In^{115} and the observations of Rowland and Bloembergen[14] on the Cu–Zn system. In that case Cu was the solvent, Zn the solute. In pure Cu, the Cu^{63} resonance is reduced by cold working to 0.4 times the intensity in the well-annealed metal. An average electric field gradient of $q=1.6\times10^{22}$ cm^{-3} produced by cold working accounts for this observation. As Zn is added to Cu, the Cu^{63} intensity is further reduced until the resonance is unobservable for Zn concentrations of more than about 20%. Rowland concluded that no Cu^{63} nuclei were observed when Zn destroyed the cubic symmetry in the first or second coordination shell. Rowland estimated that a gradient $q>1.6\times10^{23}$ cm^{-3} is needed to account for the loss of Cu^{63} nuclear contributions to the NMR line for Zn nearest and second-nearest neighbors. In a recent direct measurement, Redfield[15] showed that the quadrupole frequency for second neighbors was 1980 kc/sec (corresponding to $q=7\times10^{23}$ cm^{-3}), and for first was probably greater than 5 Mc/sec, consistent with Rowland's estimate.

In order to contrast the Pb-In results with the Cu-Zn work, it may be meaningful to think of In as the solvent,

[13] L. Schwartz and J. B. Cohen (private communication).
[13] J. G. Rider and P. L. Roney, J. Inst. Metals 91, 328 (1963).

[14] T. J. Rowland, Ph.D. thesis, Harvard University, 1954 (unpublished); N. Bloembergen and T. J. Rowland, Acta. Met. 1, 731 (1953); T. J. Rowland, Phys. Rev. 119, 900 (1960).
[15] A. G. Redfield, Phys. Rev. 130, 589 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Average electric field gradients at nuclei in metals.

	Cold worked Cu	Cu Having Zn 2*d* neighbors	In in In-Pb	Al in Al-Mg or Al-Zn alloys
$q(cm^{-3})$	1.6×10^{22}	70×10^{22}	3×10^{22}	17×10^{22}

Pb as the solute. We do not observe the In in the dilute region (high In concentration) due to the tetragonal structure. But, consider a hypothetical cold-worked cubic In metal in analogy to the cold-worked cubic Cu metal. As the solute (Pb) is added, the In^{115} intensity per atom would be expected to decrease. Thus, for example, we would expect the In^{115} resonance in the 50% In–50% Pb alloy to be unobservable (unless there is substantial clustering or ordering, which does not seem to be the case). In fact, the resonance intensity in the 50% In–50% Pb alloy is very large, corresponding to the observation of the central transition of *all* the In^{115} nuclei. Using a method described by Rowland,[16] and assuming we have no more second-order effects than discussed in Sec. II of this paper, we obtain $q<4.6\times10^{22}$ cm^{-3} for the average electric field gradient at a In^{115} nucleus for any configuration of Pb neighbors. From the fact that the satellites do not contribute to the observed lines, we find $q>1.4\times10^{22}$ cm^{-3}. Thus we can take $q\approx3\times10^{22}$ cm^{-3} as the average gradient at the In nuclei, which is comparable to the field gradient in pure cold-worked Cu, and is 20 times smaller than that found for Cu having Zn second-nearest neighbors, as is shown in Table I.

Kohn and Vosko[17] have explained the magnitude of the field gradients at the Cu site due to various solute atoms[14] on the basis of a redistribution of the conduction electron charge density near the solute.[18] The reduction in Cu NMR absorption intensity with increasing solute concentration was taken as confirmation of a long-range oscillatory behavior of the electron density. This same type of oscillatory behavior was shown to be responsible for the change in Knight shifts in silver,[19,20] cadmium,[20] and the alkali metals.[21]

In the case of Pb-In, it has already been pointed out[10] that the changes in Knight shift of the Pb^{207} upon the addition of In are small. No definite conclusion could be reached as to the importance of the charge oscillations in the lead-indium system. In any case size effects are probably too big to be neglected. We have not yet made a quantitative estimate of the field gradients. A possible explanation for both the small change in Pb^{207} Knight shift and the absence of In^{115} pronounced quadrupole effects is that the charge oscillations are greatly damped or that they oscillate rapidly and are effectively averaged out. It might then be appropriate to say that the In atoms in the cubic Pb phase screen themselves to "look almost like" Pb atoms, as far as Knight shifts and quadrupole effects are concerned. Of course, the anisotropic Knight shift broadening[10] depends on the In disrupting the cubic symmetry of the lattice so this description is not completely adequate.

It is also interesting to compare the Pb-In work with Webb's measurements[22] on Al-Zn and Al-Mg. For both Pb^{207} and Al^{27} only small changes in the Knight shift are observed. The average electric field gradient q at the site of an aluminum nucleus which is nearest neighbor of a solute (Zn or Mg) atom is found by Webb to be 17×10^{22} cm^{-3}. This is an intermediate value between In-Pb and Cu-Zn. It appears as if the effective valence difference between adjacent metals having higher group (or atomic) valences is less than for those with lower valences.

The angular correlation of the γ-γ cascade of In^{111} in thin layers has been measured[23] as a function of In concentration. This experiment shows a relatively large increase in anisotropy in going from almost pure lead to Pb-1.5% In. Kaiser attributes this increase to the electric field gradients at an indium atom due to its nearness to other indium atoms. He then compares his result to Bloembergen and Rowland's observation of the rapid loss in Cu NMR intensity upon addition of Zn. Our results show that the In-Pb result should not be compared to the Cu-Zn work. However, it is possible that an increased quadrupole interaction, reflected in the anisotropy, is associated with a first-order splitting of the satellites.

ACKNOWLEDGMENTS

We are indebted to the Consolidated Mining and Smelting Company of Canada, Limited, for the preparation of the lead-indium alloy powders. We are very grateful to L. Schwartz and Professor J. B. Cohen of Northwestern University for the search for short-range order. We thank Dr. R. M. Cotts for useful discussions.

[16] T. J. Rowland, *Progress in Material Sciences* (Pergamon Press, New York, 1961), Vol. 9.
[17] W. Kohn and S. H. Vosko, Phys. Rev. **119**, 912 (1960).
[18] See also J. Friedel, Phil. Mag. **43**, 153 (1952); A. Blandin, E. Daniel, and J. Friedel, *ibid.* **4**, 180 (1959); A. Blandin and E. Daniel, Phys. Chem. Solids **10**, 126 (1959); and A. Blandin and J. Friedel, J. Phys. Radium **21**, 689 (1960).
[19] T. J. Rowland, Phys. Rev. **125**, 459 (1962).
[20] L. E. Drain, Phil. Mag. **4**, 484 (1959).
[21] L. Rimai and N. Bloembergen, Phys. Chem. Solids **13**, 257 (1960).
[22] M. B. Webb, Phys. Chem. Solids **20**, 127 (1961).
[23] H. Kaiser, Ann. Physik **9**, 155 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Nuclear Magnetic Resonance in Lead Alloys*

Rex J. Snodgrass† and Lawrence H. Bennett
Metallurgy Division, National Bureau of Standards and Department of Physics and Astronomy, University of Maryland, College Park, Maryland

(Received 31 January 1964)

The Pb^{207} nuclear magnetic resonance has been measured in a series of lead base primary solid solutions containing Cd, In, Sn, Sb, Hg, Tl, and Bi as solutes. The central resonance frequency of Pb^{207} shifts in proportion to the solute concentration and in the opposite direction to be explained by impurity charge screening of the Thomas-Fermi type. The fractional change in Knight shift for all solutes is small $\Delta k/k \approx 0.05\ c$. The important parameter in determining the change in Knight shift seems to be the difference in valence between the impurity atoms and the lead atoms. The absorption linewidths broaden rapidly upon the addition of solute. In the absence of a quantitative calculation it is not possible to determine if the small shifts, large widths, and small quadrupole effects of In^{115} are consistent with the theory of long-range oscillations of electronic charge. However, there is evidence that this theory may be capable of explaining our data.

THE nuclear magnetic resonance (NMR) of Pb^{207} has been accurately measured in a series of lead base primary solid solutions containing second (Cd, In, Sn, Sb) and third (Hg, Tl, Bi) long period solutes. This is the first report of comprehensive Knight shift measurements on a polyvalent nontransition metal matrix. It is especially interesting to see if the theory[1] developed by Friedel, Blandin, and Daniel concerning the long-range oscillation of electronic charge density around impurity atoms in monovalent and divalent metals is capable of explaining our data for a polyvalent material. The major conclusion is that, in qualitative agreement with the Friedel theory, the valence of the solute is the controlling factor in the change of Knight shift with solute concentration.

Figure 1 shows the change in Pb^{207} resonance frequency from pure lead metal at 7 kOe and 300°K as a function of solute concentration.[2] The central resonance frequency of Pb^{207} shifts in proportion to the solute concentration (c). The shifts are in the opposite direction to be explained by impurity charge screening of the Thomas–Fermi type. The fractional change in Knight shift for all solutes is small $\Delta k/k \approx 0.05\ c$. The experimental points were omitted for clarity. The experimental error is of the order of the separation of the In and Tl lines in Fig. 1.

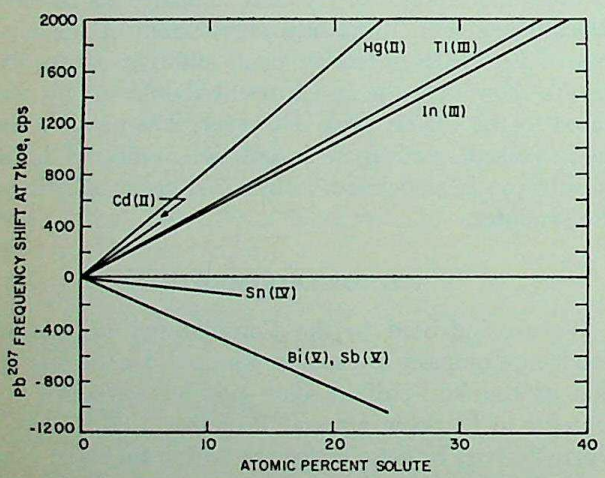


FIG. 1. Change in Pb^{207} Knight shift versus solute concentration at 300°K.

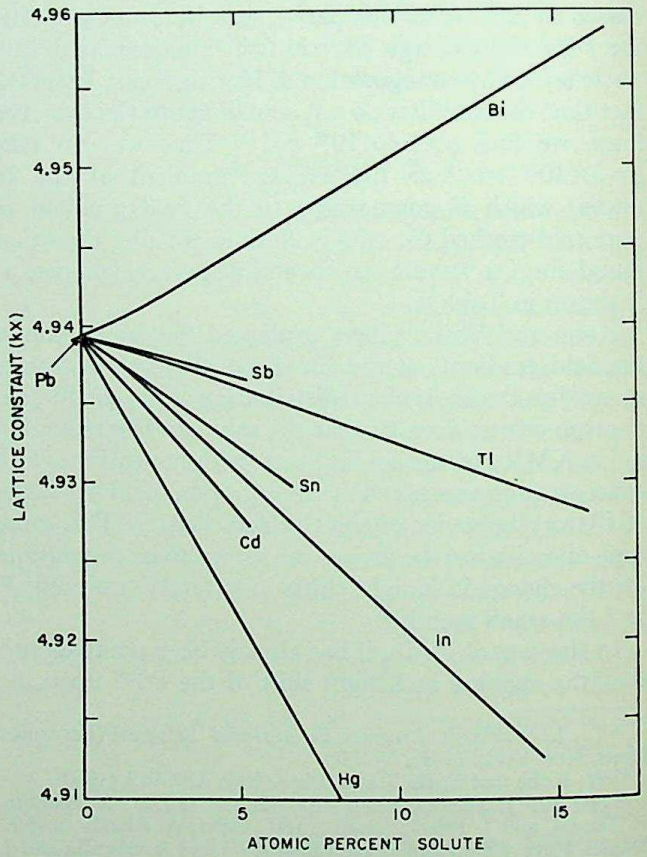


FIG. 2. Lead solid-solution lattice spacings at 20°C. After Tyzack and Raynor (Ref. 3).

* Some of the material contained herein is given in a Ph.D. thesis submitted by R. J. Snodgrass to the University of Maryland.

† 1963–1964 address: University of Paris, Faculte des Sciences-Orsay (Group II), Service de Physique des Solides, Orsay (Seine-et-Oise), France.

[1] J. Friedel, Phil. Mag. 43, 153 (1952); A. Blandin, E. Daniel, and J. Friedel, *ibid.* 4, 180 (1959); A. Blandin and E. Daniel, Phys. Chem. Solids 10, 126 (1959).

[2] The Pb^{207}-In data was described in detail in R. J. Snodgrass and L. H. Bennett, Phys. Rev. 132, 1465 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The lead solid solution lattice spacings[3] for the same solutes are shown in Fig. 2. A comparison of Figs. 1 and 2 shows that the frequency shifts partially correlate with lattice parameter changes. However, Bi and Sb give identical changes in Knight shift while the former expands and latter contracts the Pb lattice. Tl and In are more widely separated in lattice spacings than in resonance frequencies. Sn causes a larger change in lattice parameter than Tl, but a much smaller change in Knight shift (and in the opposite direction) to Tl. The correlation of the Knight shift results with Pauling or Goldschmidt radii is not good at all. From this we conclude that "size effects" are not the controlling factor in determining the Knight shift in these alloys. It is not surprising that there is partial correlation with lattice parameter data since the lattice spacing reflects not only size effects, i.e., misfit between solvent and solute, but also valence effects. The important parameter seems to be the difference in valence between the impurity atoms and the lead atoms. However, superposed on this may be a slight period dependence associated with impurities of equal valence which belong to different periods of the periodic table.[4] A pure valence effect would perhaps predict zero shift for tin. The observed shift is indeed small. Similarly, the small deviation of Cd and In from Hg and Tl, respectively, may be a period effect. However, Sb does not deviate from Bi.

The absorption linewidths are shown in Fig. 3. They broaden rapidly upon the addition of solute. It was shown previously[2] that the Pb^{207} resonance depends linearly upon the applied magnetic field for the Pb-In alloys. Since the line shapes are predominantly Lorentzian, especially for the dilute alloys, it is difficult to estimate second moments for comparison with theory. The anomalously large breadth due to Bi may be associated with its large magnetic moment, high-spin and high-isotopic abundance, together with a large exchange or pseudoexchange coupling constants. It is also possible that T_1 effects enter.

In conclusion, it seems somewhat surprising that the change in Pb Knight shift is so small, while the increase in widths is large. We believe the small change in Knight shift and the absence of pronounced quadrupole effects[5] at the quadrupolar In^{115} nucleus in Pb-In alloys are related. In the absence of a quantitative calculation it is not possible to determine if the small shifts, large widths, and small quadrupole effects of In^{115} are consistent with the theory of long-range oscillations of electronic charge. However, there is evidence that this theory may be capable of explaining our data. (1) The shifts correlate with valence. (2) The sign of the shifts is in the opposite direction to be explained by Thomas-Fermi screening. (3) The absorption linewidths are field-dependent. This might also be due to anisotropic Knight-shift broadening, but we have not excluded inhomogeneous isotropic Knight-shift broadening.

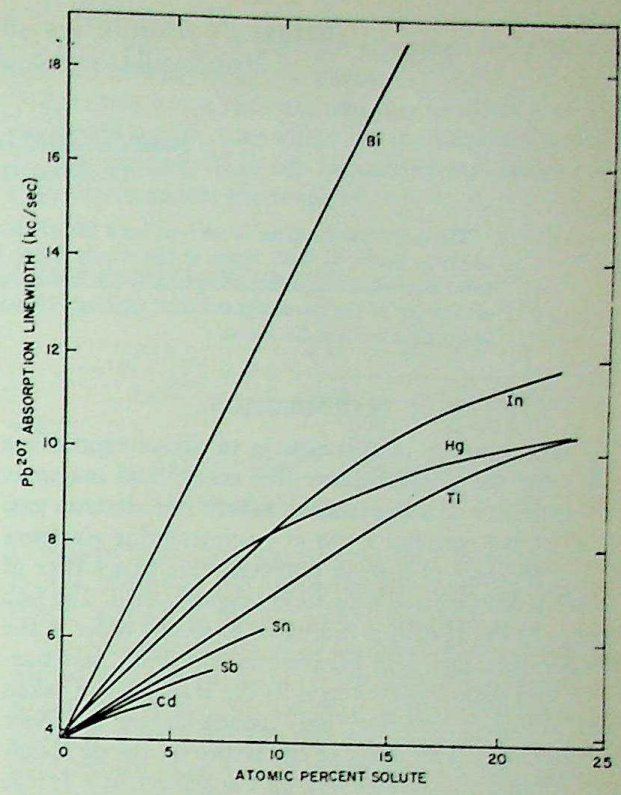


FIG. 3. Pb^{207} absorption linewidths versus solute concentration at 300°K and 7 kOe.

ACKNOWLEDGMENTS

We are very grateful to D. P. Fickle for preparing the samples and to C. J. Bechtoldt for x-ray measurements.

[3] C. Tyzack and G. V. Raynor, Acta Cryst. 7, 505 (1954).
[4] F. J. Blatt, Phys. Rev. 108, 285 (1957).
[5] L. H. Bennett and R. J. Snodgrass, Bull. Am. Phys. Soc. 7, 228 (1962), and to be published.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Magnetic Properties of Nearly Free Electrons.* Nonoscillatory Magnetic Susceptibility

M. L. GLASSER
Battelle Memorial Institute, Columbus, Ohio
(Received 30 December 1963)

The partition function is derived for a gas of electrons in the presence of a weak periodic potential and uniform magnetic field. From it the steady terms in the free energy and the zero-field–zero-temperature total electronic magnetic susceptibility are calculated. The results are applied to the alkali metals. The nonadditivity of the paramagnetic and diamagnetic susceptibilities and the inadequacy of the effective mass approximation are discussed.

I. INTRODUCTION

THE purpose of this note is to derive expressions for the nonoscillatory free energy and magnetic susceptibility of a degenerate, nearly free electron gas. That is, we consider a gas of noninteracting electrons in the presence of a weak potential, due to a lattice of positive charges, and a uniform magnetic field. The free energy of the system is evaluated to second order in the temperature and field parameters with both paramagnetic and diamagnetic terms in the Hamiltonian taken into account; however, only nonoscillatory behavior is considered. The effect of the lattice on the de Haas–van Alphen oscillations will be the subject of a future report.

The calculation was made with two ends in view. First, considering the success of the pseudopotential concept,[1–3] it appears that the nearly free electron approximation has computational value as well as being merely instructive. Second, we wish to go beyond the limitations of the effective mass approximation for describing lattice effects in the context of magnetic properties. A calculation is performed in Sec. III using available pseudopotential parameters for the alkali metals Li, Na, and K. There exists no completely satisfactory calculation of the role of correlation in metallic properties for the density range into which these metals fall and, since this topic is outside the scope of the present study, we shall only point out that corrections can be estimated by interpolation methods.[4] Quantitative accord with experiment thus can not be expected, but it is proposed that for metals with small energy gaps the effect of the lattice is qualitatively accounted for. The criterion suggested is that the ratio of the energy gaps to the Fermi level be much less than unity.

II. CALCULATION

We consider N noninteracting electrons moving in a weak periodic potential $V(r)$, which occupies a very large volume Ω, in the presence of a magnetic field represented by the vector potential $\mathbf{A}=(-H_0y,0,0)$. An electron will therefore be described by the Hamiltonian

$$\mathcal{H}=(1/2m)(\mathbf{P}-(e/c)\mathbf{A})^2+2\mu_0H_0S_z+V(\mathbf{r})\equiv\mathcal{H}_0+V(\mathbf{r}). \quad (1)$$

As is well known,[5] the eigenstates of $\mathcal{H}_0$ are (apart from normalization)

$$|k_xnk_z,\sigma\rangle=e^{-\frac{1}{2}(\eta-\eta_0)^2}H_n(y-y_0)e^{i(k_zx+k_zz)}\chi_\sigma, \quad (2a)$$

$$E_\sigma(n,k_z)=2\mu_0H_0(n+1/2)+\hbar^2k_z^2/2m+\mu_0H_0\sigma, \quad (2b)$$

where μ_0=Bohr magneton; χ_σ=spin state, $\sigma=\pm1$; $\eta=(eH/\hbar c)^{1/2}y$, $\eta_0=-(\hbar c/eH)^{1/2}k_x$; H_n=nth Hermite polynomial.

The thermodynamic properties of the system can be studied conveniently by means of the partition function

$$Z(\gamma)=\mathrm{Tr}\{\exp[-\gamma(\mathcal{H}_0+V)]\}, \quad \gamma=1/kT. \quad (3)$$

Using the Schwinger trace formula[6] (the variable of integration has been changed to $u=\gamma s_1$)

$$Z(\gamma)\cong\mathrm{Tr}\{e^{-\gamma\mathcal{H}_0}\}-\gamma\,\mathrm{Tr}\{V(\mathbf{r})e^{-\gamma\mathcal{H}_0}\}+\tfrac{1}{2}\gamma^2\,\mathrm{Tr}\left\{\int_0^1 du\,V(\mathbf{r})e^{-\gamma\mathcal{H}_0(1-u)}V(\mathbf{r})e^{-\gamma\mathcal{H}_0u}\right\}\equiv Z_0(\gamma)-Z_1(\gamma)+Z_2(\gamma). \quad (4)$$

Since the Zeeman term commutes with the remainder of the Hamiltonian, the spin part of the trace is performed very simply to give

$$Z(\gamma)=2\cosh(\mu_0H_0\gamma)[z_0(\gamma)-z_1(\gamma)+z_2(\gamma)]. \quad (5)$$

$z_0(\gamma)$ is known from the work of Sondheimer and

---

* This study was supported in part by U. S. Air Force Office of Scientific Research Grant No. AF-AFOSR-260-63.

[1] J. C. Phillips and L. Kleinman, Phys. Rev. 116, 287 (1959).
[2] M. H. Cohen and V. Heine, Phys. Rev. 122, 1821 (1961).
[3] J. M. Ziman, Phil. Mag. 6, 1013 (1961).
[4] S. D. Silverstein, Phys. Rev. 130, 1703 (1963).

[5] L. Landau and E. Lifshitz, *Quantum Mechanics*, (Addison-Wesley Publishing Company, Reading, Massachusetts, 1958), Sec. 125.

[6] A. Saenz and R. O'Rourke, Rev. Mod. Phys. 27, 381 (1955). There is a misprint in Eq. (1.2.8) which should read:

$$\mathrm{Tr}\{Q(s)\}=\mathrm{Tr}\{e^{-sa}\}-s\,\mathrm{Tr}\{be^{-sa}\}+s\sum_1(-1)^n/n\times\mathrm{Tr}\left\{be^{-sa}\int_0^s ds_1\cdots\int_0^{s_{n-2}}ds_{n-1}b(s_1)\cdots b(s_{n-1})\right\}.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Wilson[7]:

$$z_0(\gamma)=\Omega\alpha\gamma^{-3/2}(\mu_0 H_0\gamma)\,\mathrm{csch}(\mu_0 H_0\gamma),\qquad \alpha=(m/2\pi\hbar^2)^{3/2}. \tag{6}$$

$z_1(\gamma)$ can be calculated directly if $V(\mathbf{r})$ is represented as a Fourier series

$$V(\mathbf{r})=\sum_{\mathbf{K}} V_{\mathbf{K}} e^{-i\mathbf{K}\cdot\mathbf{r}}, \tag{7}$$

where the $\mathbf{K}$'s are reciprocal lattice vectors. Representing the trace in the form

$$\mathrm{Tr}\{B\}=\frac{\Omega^{2/3}}{(2\pi)^2}\int dk_x\int dk_z\sum_{n=0}^{\infty}\langle k_x n k_z|B|k_x n k_z\rangle, \tag{8}$$

and using the generating function for products of Hermite polynomials[8] we obtain

$$z_1(\gamma)=\gamma V_0 z_0(\gamma), \tag{9}$$

where V_0 is the average crystal potential. For simplicity it will be assumed that V has a center of symmetry.

The evaluation of $z_2(\gamma)$ is straightforward but tedious. We begin with the expression

$$z_2(\gamma)=\tfrac{1}{2}\gamma^2\frac{\Omega^{4/3}}{(2\pi)^4}\int_0^1 du\int dk_x dk_x'\int dk_z dk_z'\times\sum_{n,n'}|\langle k_x n k_z|V(r)|k_x' n' k_z'\rangle|^2\times\exp\{-\gamma E(n',k_z')(1-u)+E(n,k_z)u\}. \tag{10}$$

First the spatial integrals over $\mathbf{r}$ and $\mathbf{r}'$ in the matrix elements can be done, except for the y and y' integrals, making use of (8). The sums over n and n' are evaluated next using the generating function[8] and the y and y' integrals then performed in terms of the variables $y\pm y'$. At this point the remaining integrals, except for the u integration, are done without difficulty leading to the result

$$z_2(\gamma)=\tfrac{1}{2}\gamma^2 z_0(\gamma)\sum_{\mathbf{K}} V_{\mathbf{K}}^2\times\int_0^1\exp\left\{-\frac{\hbar c}{eH}\mathbf{K}_\perp^2\mu-\frac{\gamma\hbar^2}{2m}\mathbf{K}_{\parallel}^2 u(1-u)\right\}du. \tag{11}$$

$\mathbf{K}_\perp$ and $\mathbf{K}_{\parallel}$ are the components of $\mathbf{K}$, perpendicular and parallel to the magnetic field, and

$$\mu^{-1}=\coth(\mu_0 H_0\gamma u)+\coth[\mu_0 H_0\gamma(1-u)]. \tag{12}$$

Combining (5)–(11) we have

$$Z(\gamma)=Z_0(\gamma)\{1-\gamma V_0+\tfrac{1}{2}\gamma^2 V_0^2+\tfrac{1}{2}\gamma^2\sum{}' G(\mathbf{K},\gamma)V_{\mathbf{K}}^2\}, \tag{13}$$

where $G(\mathbf{K},\gamma)$ is the integral in (11) and

$$Z_0(\gamma)=2\Omega\alpha\gamma^{-3/2}(\mu_0 H_0\gamma)\coth(\mu_0 H\gamma) \tag{14}$$

is the partition function for free electrons. For convenience, the term $K=0$ has been separated from the sum and this is denoted by the prime.

For comparison the partition function has also been calculated for the case where the diamagnetic terms in (1) are omitted. This can be carried out easily using plane-wave states; the result is

$$Z'(\gamma)=Z_0'(\gamma)\left\{1-\gamma V_0+\tfrac{1}{2}\gamma^2 V_0^2+\gamma^{5/2}\binom{1}{-\alpha}\sum{}' V_K^2\int\frac{d^3k}{(2\pi)^3}\frac{e^{-\gamma\epsilon(k)}}{\epsilon(k-K)-\epsilon(k)}\right\}, \tag{13'}$$

where

$$Z_0'(\gamma)=2\Omega\alpha\gamma^{-3/2}\cosh(\mu_0 H_0\gamma),\qquad \epsilon(k)=\hbar^2k^2/2m.$$

It follows from these expressions that to second order in the potential the paramagnetic and diamagnetic effects of the field are not simply additive.

The free energy and other thermodynamic properties can now be calculated. Since we consider here only the steady susceptibility, we may simply disregard the poles of (13) lying off the real axis in the inverse temperature plane.[9] It is convenient at this point to separate $Z(\gamma)$ into field-dependent and independent parts retaining terms of no higher than second order in the field. Introducing the parameter $\beta=\mu_0 H_0$ we find $\mu\cong\beta\gamma S[1-\frac{1}{3}(\beta\gamma)^2 S]$ where $S=u(1-u)$ and $G(\mathbf{K},\gamma)=\frac{1}{2}I_0[\frac{1}{4}\gamma\epsilon(\mathbf{K})]+\beta^2\gamma^3(\epsilon(K_\perp)/96)I_2[\frac{1}{4}\gamma\epsilon(\mathbf{K})]$, where

$$I_n(x)=\int_0^1\frac{y^n e^{-xy}}{(1-y)^{1/2}}. \tag{15}$$

Thus,

$$Z(\gamma)=Z^{(0)}(\gamma)+\beta^2 Z^{(1)}(\gamma),$$

where

$$Z^{(0)}(\gamma)=2\Omega\alpha\gamma^{-3/2}\{1-\gamma V_0+\tfrac{1}{2}\gamma^2 V_0^2+\tfrac{1}{4}\gamma^2\sum{}' V_K^2 I_0[\tfrac{1}{4}\gamma\epsilon(K)]\}, \tag{16a}$$

$$Z^{(1)}(\gamma)=2\Omega\alpha\gamma^{-3/2}\{\tfrac{1}{3}\gamma^2-\tfrac{1}{3}\gamma^3 V_0+\tfrac{1}{6}\gamma^4 V_0^2+\tfrac{1}{12}\gamma^4\sum{}' V_K^2 I_0[\tfrac{1}{4}\gamma\epsilon(K)]\}+(1/192)\gamma^5\sum{}' V_K^2\epsilon(\mathbf{K}_\perp)I_2[\tfrac{1}{4}\gamma\epsilon(K)]. \tag{16b}$$

The quantity $\Phi=F-n\zeta$, $(n=N/\Omega)$, will now be calculated using Eq. (14) of Ref. 9. Asymptotic expansions in powers of the degeneracy parameter $(\gamma\zeta)$ can be obtained for the contributions of the first three terms of (16a) and (16b) to obtain Φ. For the remaining terms the procedure described in Ref. 9 would be permitted only if $\gamma\zeta\gg\frac{1}{4}\gamma\epsilon(K)$; since this is not the case for the alkali metals, the relevant parameter for these terms is $a=4\zeta/\epsilon(K)$.

The calculation as outlined gives

$$\Phi=\Phi_0^{(0)}+\gamma^{-2}\Phi_1^{(0)}+\beta^2\Phi_0^{(1)}+(\beta/\gamma)^2\Phi_1^{(1)}, \tag{17}$$

[7] E. H. Sondheimer and A. H. Wilson, Proc. Roy. Soc. (London) **A210**, 173 (1951).
[8] Eq. (A4) of Ref. 7.
[9] M. L. Glasser, J. Math. Phys. (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$\Phi_0^{(0)} = -(2/\sqrt{\pi})\alpha\{(8/15)\zeta^{5/2} - \tfrac{4}{3}V_0\zeta^{3/2} + V_0^2\zeta^{1/2} + \tfrac{1}{4}\sum' V_K^2[\epsilon(K)]^{1/2} f(a)\}, \quad (18a)$$

$$\Phi_1^{(0)} = -\tfrac{1}{3}\pi^2(2\alpha/\sqrt{\pi})\{\zeta^{1/2} - \tfrac{1}{2}V_0\zeta^{-1/2} - \tfrac{1}{8}V_0^2\zeta^{-3/2} + \sum' V_K^2[\epsilon(K)]^{-3/2} g(a)\}, \quad (18b)$$

$$f(a) = \sqrt{a} + (a-1)\tanh^{-1}\sqrt{a},$$

$$g(a) = a^{-1/2}(1-a)^{-1}; \quad (19)$$

$$\Phi_0^{(1)} = -(2\alpha/\sqrt{\pi})\{\tfrac{2}{3}\zeta^{1/2} - \tfrac{1}{3}V_0\zeta^{-1/2} - \tfrac{1}{12}V_0^2\zeta^{-3/2} + \tfrac{2}{3}\sum' V_k^2[\epsilon(K)]^{-3/2} g(a) + \tfrac{1}{6}\sum' V_K^2(\epsilon(\mathbf{K}_1)/[\epsilon(K)]^{5/2}) h(a)\}, \quad (20a)$$

$$\Phi_1^{(1)} = -\tfrac{1}{3}\pi^2(2\alpha/\sqrt{\pi})\{-\tfrac{1}{12}\zeta^{-3/2} - \tfrac{1}{8}V_0\zeta^{-5/2} - \tfrac{5}{32}V_0^2\zeta^{-7/2} + (16/3)\sum' V_K^2[\epsilon(K)]^{-7/2} g''(a) + \tfrac{4}{3}\sum' V_K^2(\epsilon(\mathbf{K}_1)/[\epsilon(K)]^{9/2}) h''(a)\}, \quad (20b)$$

and

$$h(a) = \tfrac{3}{2}\tanh^{-1}\sqrt{a} + \frac{a^{1/2}(5-3a)}{2(1-a)^2}. \quad (21)$$

In order to use these expressions, the chemical potential must be known to second order in the various parameters. It can be calculated by solving the equation $(\partial\Phi/\partial\zeta)_\Omega + n = 0$. Except in rare cases, the temperature and field dependences of ζ are small enough to be neglected. Hence we need retain only $\Phi_0^{(0)}$ in this equation and solve for ζ to second order in the lattice potential. This calculation is easily carried out and yields

$$\zeta = \zeta_0[1 + (V_0/\zeta_0) - \tfrac{1}{2}\zeta_0^{3/2}\sum' V_K^2[\epsilon(K)]^{-1/2} f'(a_0)], \quad (22)$$

where $\zeta_0 = \frac{1}{4}(9\pi n^2)^{1/3}(2\pi\hbar^2/m)$ and $a_0 = 4\zeta_0/\epsilon(K)$.

The magnetic susceptibility can now be calculated. Putting (17) into the formula $\chi = -(1/H)(\partial F/\partial H)|_\Omega$ we find $\chi = -2\mu_0^2[\Phi_0^{(1)} + (kT)^2\Phi_1^{(1)}]$. To second order, the temperature-independent susceptibility is

$$\chi_0 = \chi_0^{\text{free}}[1 - (1/8\zeta_0^2)\sum' V_K^2 F(a_0)], \quad (23)$$

where

$$F(a_0) = (a_0)^{1/2}\tanh^{-1}(a_0)^{1/2} - \frac{a_0}{1-a_0} - \frac{1}{4}\frac{\epsilon(\mathbf{K}_1)}{\epsilon(K)}a_0^{3/2}h(a_0), \quad (24)$$

and $\chi_0^{\text{free}} = \mu_0^2/\zeta_0$ is the susceptibility for free electrons. A similar formula can be obtained for the temperature-dependent term.

An equation, resembling (23), for the paramagnetic susceptibility was published recently by Abe.[10] His formula is derived along the lines of this calculation in the Appendix.

III. DISCUSSION

In this section the results of the preceding analysis are applied to the alkalis Li and Na, these being the only metals for which the spin susceptibility has been measured. We shall cast our analysis for these metals in the form of a one-parameter model in terms of the lowest nonzero Fourier coefficient of the crystal potential. This is determined by the energy gap at the zone surface and estimates from band calculations[3,10] are listed in Table I. An estimate for sodium based on Callaway's[11] pseudopotential showed that the sum in (23) is given almost entirely by the contributions fron the reciprocal lattice vectors of shortest length. The results obtained by summing over only these vectors and using the Fourier coefficients listed in Table I are shown in Table II. $F(a_0)$ is a universal function for a given lattice; for the bcc case $F(a_0) = -6.33$ for the vectors of type $2\pi/a$ (1, −1, 0) and −4.29 for those of type $2\pi/a$ (1,0,1) where a is the lattice constant. The calculation for Na carried out using Callaway's pseudopotential gave a result much larger than that listed in Table II. This is consistent with the observation that the energy gap (at the point N) found by Callaway is an order of magnitude larger than that found in most other calculations.

Estimates of the correction to the susceptibility due to correlation have been made on the basis of Silverstein's[4] work and are listed under χ^* in Table I. The entries in Table II denoted by a star were obtained by including these adjustments. The starred values of the paramagnetic susceptibility are in close agreement with Abe's[10] values, obtained by a Sampson-Seitz procedure. Considering that the sum in (23) has been underestimated and that correlation has been accounted for in a very crude fashion, agreement with experiment for Li and Na appears to be satisfactory.

Also listed in Table II are values for the "diamagnetic susceptibility" obtained by subtracting χ_p (A5) from

TABLE I. Pseudopotential and free electron parameters for the alkali metals.

	ζ_0 (Ryd)	V_{K_0} (Ryd)	χ_0^{free} [a]	χ^* [a]
Li	0.355	0.112	0.54	0.25
Na	0.235	0.010	0.43	0.20
K	0.155	−0.018	0.35	0.18

[a] ($\times 10^6$ cgs vol. units).

TABLE II. Experimental and calculated susceptibility ($\times 10^6$ cgs vol. units).

	Li	Na	K
χ_0 [a]	1.20*	0.64*	0.56*
	0.95	0.44	0.38
χ_0 [b,c]	1.94±0.05	0.88±0.03	
χ_p [d]	1.33*	0.84*	0.73*
	1.08	0.66	0.55
χ_p [b]	2.08±0.10	0.95±0.10	
χ_d (calc)	−0.13	−0.22	−0.17
χ_d [b]	−0.14±0.15	−0.07±0.11	

[a] From Eq. (23).
[b] R. T. Schumacher and C. P. Slichter, Phys. Rev. 101, 58 (1956).
[c] The ionic diamagnetism has been subtracted.
[d] From Eq. (A5).

[10] R. Abe, Progr. Theoret. Phys. (Kyoto) 29, 23 (1963).

[11] J. Callaway, Phys. Rev. 112, 322 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

χ_0 (23). Because of the nonadditivity of diamagnetic and paramagnetic effects, discussed below in more detail, these values correspond only roughly to the Landau-Peierls diamagnetism. There is some evidence[12] that the diamagnetic susceptibility is only slightly affected by correlation; thus, much of the correlation contribution is subtracted off by this procedure. The close agreement between theory and experiment here is strong evidence in favor of a pseudopotential theory of these metals.

In the course of this calculation it was discovered that the same basic method has been applied by Samoilovich and Rabinovich[13] to study the diamagnetism of nearly free electrons; where comparable, their results agree with this calculation. However, due to the nonadditivity of the susceptibilities, the result of adding their expression to Abe's differs from (23); indeed

$$\Delta=[\chi_d(\mathrm{SR})+\chi_p(\mathrm{Abe})-\chi_0(23)]/\chi_0{}^{\mathrm{free}}$$
$$=\frac{1}{16\zeta_0{}^2}\sum{}'\, V_K{}^2\frac{\epsilon(\mathbf{K}_1)}{\epsilon(K)}[a_0(1-a_0)^{-2}-\tfrac{1}{2}a_0{}^{3/2}h(a_0)].$$

For lithium, $\Delta=0.62$, while for sodium, $\Delta=0.012$. Therefore care must be taken in separating the paramagnetic and diamagnetic susceptibilities for comparison with experiment. Since the diamagnetic susceptibility is not directly measured, it is believed that the estimates in Table II are more significant than those of Samoilovich and Rabinovich. The existence of this nonadditivity has been discussed in a general way by Blount.[14] This feature would appear to be a source of considerable error in previous calculations of diamagnetic susceptibilities.

In conclusion, a comparison can be made with the effective-mass approximation within the scope of the nearly free electron model. The susceptibility is given in terms of the effective mass ratio by[7]

$$\chi_0/\chi_0{}^{\mathrm{free}}=\tfrac{3}{2}(m^*/m)^{3/2}-\tfrac{1}{2}(m/m^*)^{1/2}. \quad (25)$$

In the nearly free electron approximation m^*/m, averaged over the Fermi surface, is

$$m^*/m=1+(1/4\zeta_0{}^2)\sum{}'\, V_K{}^2(a_0)^{1/2}\tanh^{-1}(a_0)^{1/2}. \quad (26)$$

Combining (25) and (26) gives

$$\chi_0/\chi_0{}^{\mathrm{free}}=1+(5/8\zeta_0{}^2)\sum{}'\, V_K{}^2(a_0)^{1/2}\tanh^{-1}(a_0)^{1/2},$$

which differs significantly from (23).

ACKNOWLEDGMENTS

The author is grateful to Dr. V. E. Wood and Dr. F. J. Milford for conversations on this topic.

[12] H. Kanazawa and N. Matsudaira, Progr. Theoret. Phys. (Kyoto) 23, 433 (1960).
[13] A. Samoilovich and E. Rabinovich, Fiz. Tverd. Tela 5, 778 (1963) [English transl.: Soviet Phys.—Solid State 5, 567 (1963)].
[14] E. I. Blount, Phys. Rev. 126, 1636 (1962).

APPENDIX

Using (13′) and proceeding as in the body of the text we find

$$\Phi'=\Phi_0'+\Phi_1'+\Phi_2'+\Phi_2'', \quad (A1)$$

where

$$\Phi_0'=-\pi/\gamma(m/2\pi\hbar^2\gamma)^{3/2}[B_5\{\gamma(\zeta+\mu_0H)\}+B_5\{\gamma(\zeta-\mu_0H)\}],$$
$$\Phi_1'=\pi V_0(m/2\pi\hbar^2\gamma)^{3/2}[B_3\{\gamma(\zeta+\mu_0H)\}+B_3\{\gamma(\zeta-\mu_0H)\}],$$
$$\Phi_2'=-\pi\gamma/2V_0{}^2(m/2\pi\hbar^2\gamma)^{3/2}[B_1\{\gamma(\zeta+\mu_0H)\}+B_1\{\gamma(\zeta-\mu_0H)\}],$$

and

$$B_\sigma(\lambda)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}s^{-\sigma/2}\csc\pi s e^{\lambda s}ds$$

is an integral discussed in another paper.[9] It is convenient to handle Φ_2'', corresponding to the last term in (13′), somewhat differently. We first calculate $z''(E)$, the inverse Laplace transform of the corresponding term in $Z'(\gamma)/\gamma^2$. Then we obtain

$$\Phi_2''=\sum{}'\, V_K{}^2\int_0^\infty\frac{\partial f_0}{\partial E}\{g(K,E+\mu_0H)+g(K,E-\mu_0H)\}dE,$$

where f_0 is the Fermi function and

$$g(K,\lambda)=(2\pi)^{-3}\int_{\epsilon(k)<\lambda}d^3k[\epsilon(\mathbf{k}-\mathbf{K})-\epsilon(\mathbf{k})]^{-1}.$$

In the zero-temperature limit (A1) becomes

$$-\Phi(T=0)=[1/\Gamma(\tfrac{7}{2})](m/2\pi\hbar^2)^{3/2}\times\{[(\zeta+\mu_0H)^{5/2}+(\zeta-\mu_0H)^{5/2}]-\tfrac{5}{2}V_0[(\zeta_0+\mu_0H)^{3/2}+(\zeta-\mu_0H)^{3/2}]+(15/8)V_0{}^2[(\zeta+\mu_0H)^{1/2}+(\zeta-\mu_0H)^{1/2}]\}+\sum{}'\, V_K{}^2\{g(K,\zeta+\mu_0H)+g(K,\zeta-\mu_0H)\}. \quad (A2)$$

This leads to

$$\chi_p=\frac{4}{\sqrt{\pi}}\mu_0{}^2\left(\frac{m}{2\pi\hbar^2}\right)^{3/2}\zeta^{1/2}\left[1-V_0/\zeta-\tfrac{1}{8}(V_0/\zeta)^2+\frac{2\pi}{\zeta}\left(\frac{m}{2\pi\hbar^2}\right)\sum{}'\frac{V_k{}^2}{K^2-4k_0{}^2}\right], \quad (A3)$$

where k_0 is the free electron Fermi momentum. The chemical potential is obtained as before. Inserting (22) into (A3) leads to Abe's[10] expression

$$\chi_p/\chi_p{}^{\mathrm{free}}=1+(1/8\zeta_0{}^2)\times\sum\frac{V_K{}^2}{u^2-1}\left\{1+\frac{1-u^2}{2u}\ln\left|\frac{u+1}{u-1}\right|\right\}, \quad (A4)$$

where $\chi_p{}^{\mathrm{free}}=3\mu_0{}^2n/2\zeta_0$, and $u=K/2k_0=1/\sqrt{a_0}$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Electron Paramagnetic Resonance of Fe^{3+} in the Strong Axial Field of $PbTiO_3$ Host

DENIS J. A. GAINON
Electronic Research Division, Clevite Corporation, Cleveland, Ohio

(Received 16 December 1963; revised manuscript received 10 February 1964)

An electron paramagnetic resonance investigation has been made of single crystals of tetragonal $PbTiO_3$. A spectrum attributed to Fe^{3+} was found which can be described by the spin Hamiltonain: $\mathcal{H}=g_{\parallel}\beta H_z S_z + g_{\perp}\beta(H_x S_x + H_y S_y)$ with $S=\frac{1}{2}$, $g_{\parallel}=2.009\pm0.005$ and $g_{\perp}=5.97\pm0.02$ at room temperature. These values are explained in terms of a strong tetragonal component of the crystalline field.

INTRODUCTION

LEAD titanate is a ferroelectric which is isomorphous at room temperature with tetragonal barium titanate and possesses a Curie point at 490°C, where the structure changes to cubic symmetry. The cell dimensions are $a=3.904$ Å and $c=4.150$ Å at 20°C. Small single crystals can be grown, but ceramics of $PbTiO_3$ tend to fracture when cooled through the Curie point. The purpose of this work has been to study the paramagnetic resonance spectrum of $PbTiO_3$ and to correlate the results with the known properties of this substance.

MEASUREMENTS AND RESULTS

Single crystals of $PbTiO_3$, grown from a KF flux melt in platinum crucibles, occur in the form of small plates or cubes up to several millimeters which can be yellow, green, or black. It is not at present possible to grow large single domain crystals.

Measurements were made primarily at room temperature[1] using a Varian V-4502 spectrometer working at about 8.8 Gc/sec with a magnetic field modulation frequency of 400 cps.

All the crystals studied show the same spectrum, which is essentially composed of three resonance lines. The lines are relatively broad; the absorption derivative peak-to-peak value is about 50 G. Variation of the orientation of the magnetic field relative to the crystallographic axes shows that each of the lines belongs to a paramagnetic impurity of effective spin $S=\frac{1}{2}$ and follows the spin-Hamiltonian

$$\mathcal{H}=g_{\parallel}\beta H_z S_z+g_{\perp}\beta(H_x S_x+H_y S_y),$$

with

$$g_{\parallel}=2.009\pm0.005,$$
$$g_{\perp}=5.97\pm0.02.$$

The only difference between the three lines is the orientation of the Z axis which, for each line, corresponds to one of the crystal axes. It is therefore natural to conclude that the spectrum is produced by the same impurity in the three possible orientations of domains (twinning of 180° being magnetically unobservable). This is confirmed by the fact that the relative intensity of the lines is strongly variable from one crystal to another, showing a different distribution of the domains. The three lines can be roughly equal; but also one or two of them can be of weaker intensity or unobserved, particularly in the crystals occurring in the form of plates. Among the cubes, the most extreme departure from random distribution was found in a crystal possessing about 80% of the domains in the same orientation.

The identification of $g_{\perp}$ is given by rotation of the magnetic field around one of the crystal axes. In this case the domains whose Z axes always remain perpendicular to the direction of H correspond to the orientation independent resonance line.

DISCUSSION

The measurements show that the paramagnetic impurity responsible for the spectrum is located in a position of local tetragonal symmetry and therefore probably substituted for an ion in the lattice.

Several considerations allow us to identify this impurity as Fe^{3+}. The signal is relatively strong, and spectrographic analysis of nominally pure samples shows that iron is the impurity present in the highest concentration (between 0.05 and 0.5 at. %). Crystals intentionally iron-doped gave the same spectrum; no additional resonance lines appeared. The resonance of Fe^{3+} is relatively strong in the other titanium compounds of perovskite-type structure, $BaTiO_3$[2] and $SrTiO_3$,[3] which are grown basically by the same method. No trace was found of the five-line spectrum characteristic of Fe^{3+} in tetragonal symmetry in any of the $PbTiO_3$ samples. The iron could possibly be in a reduced state (Fe^{2+} is expected to be unobserved at room temperature). The oxidation of Fe^{2+} to Fe^{3+} was apparently done without difficulty in $BaTiO_3$.[2] In $PbTiO_3$, heating in air for several hours, at temperatures up to 1000°C, then in oxygen (flow of 3.25 l/min for 40 h at 800°C) did not change the spectrum, suggesting that the iron was in the trivalent state prior to the oxidation.

[1] Investigation of the spectrum up to 300°C and down to −120°C did not produce any fundamental change. In particular, the transition to a multiple cell structure reported at −100°C when cooling slowly (see Ref. 9) did not produce a transformation of the spectrum.

[2] A. W. Hornig, R. C. Rempel, and H. E. Weaver, Phys. Chem. Solids **10**, 1 (1959).

[3] K. A. Müller, Helv. Phys. Acta **31**, 173 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Although other ions which may be paramagnetic are present in the crystal, i.e., titanium, platinum,[4] niobium, and copper,[5] they are not expected to have the observed g values and spectral intensity.

Because of size considerations, Fe^{3+} is expected to substitute for Ti^{4+} in the lattice, and its paramagnetic spectrum can be explained in terms of a strong axial field. The problem was theoretically studied by Griffith.[6] Fe^{3+} in its high-spin form is in the ${}^6S_{5/2}$ state. Under the action of a tetragonal crystalline field, this ground state splits into three twofold degenerate levels and its behavior in the presence of a magnetic field can be described by an adequate spin-Hamiltonian. If, however, the coefficient D of S_z^2 in the spin-Hamiltonian, proportional to the zero-field splitting, is much larger than the magnetic field energy, the lowest doublet is not mixed with the highest ones and the spectrum is described by a new spin-Hamiltonian with $S'=\frac{1}{2}$, $g_{11}'\cong 2$, and $g_1'\cong 6$. This effect was observed by Griffith and Bennett *et al.*,[7] but with relatively low precision in a few organic compounds (haemoglobin derivatives).

Several facts help to explain the large zero-field splitting of Fe^{3+} in $PbTiO_3$. The work of the above authors suggests that a D of at least 1 cm^{-1} is required to observe a true g_1 of 6 if the minimum wavelength available is about 3 cm. Although studies at shorter wavelengths would provide further information, a minimum D of 1 cm^{-1} seems consistent with the crystalline structure of $PbTiO_3$. The relation between D and the coefficient A_2^0 of Y_2^0 in the development of the crystalline field potential is not completely understood, but an extrapolation of the tentative quadratic relation given by Nicholson and Burns[8] leads to an A_2^0 at least two times larger in $PbTiO_3$ than in $BaTiO_3$. This is in accord with the main feature in $PbTiO_3$, which is the large displacement of the Pb and Ti ions, leading to a strong tetragonal component. Pb moves through 0.47 Å (0.05 Å for Ba in $BaTiO_3$) and Ti through 0.30 Å (0.10 Å in $BaTiO_3$) relative to the almost undistorted framework of oxygen ions. The ratio c/a of the cell dimensions is 1.063 in $PbTiO_3$, about six times larger. The spontaneous polarization calculated from ionic displacements is about three times larger than for $BaTiO_3$.[9] Thus, the observed EPR spectrum of Fe^{3+} appears consistent with the strong tetragonal crystalline field of $PbTiO_3$. It should be noted that possible distortions due to the noncompensated charge of Fe^{3+} may locally modify the crystal field.

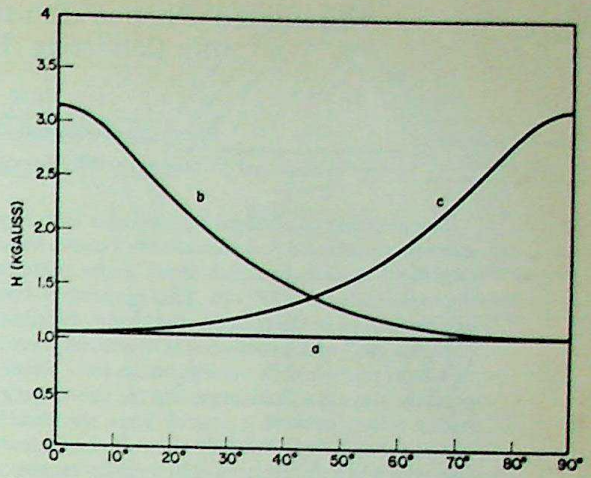


FIG. 1. Angular dependence of the three resonance lines for rotation of the magnetic field around one of the symmetry axes. Curve (a) corresponds to domains whose Z axes always remain perpendicular to the magnetic field. Curves (b) and (c) correspond to domains whose Z axes lie in the plane of rotation of H.

After this work was completed we learned from Müller[10] that similar resonances have been observed in cubic $SrTiO_3$ which have an intensity about 0.15 times that of the central transition of the cubic spectrum. Accurate determination of g_1 at two different frequencies allowed these workers to calculate a value of $|2D|=2.85\pm0.15$ cm^{-1}. They concluded that the only possible explanation for such a large D is that local charge compensation takes place at a nearest oxygen site. An interpretation of the $PbTiO_3$ spectrum in terms of local charge compensation would imply (1) that this charge compensation occurs only along the direction of polarization (since the resonance lines are of unequal intensities) and (2) that all the Fe^{3+} ions are locally compensated in an equivalent manner (since no other spectrum due to Fe^{3+} is detected).

ACKNOWLEDGMENTS

The author received crystals both from F. J. Scholz and Sandia Corporation, Albuquerque. He wishes to thank A. P. Greifer for the oxidation experiment, Dr. T. R. Sliker for valuable help and discussions, and Professor S. H. Christensen of Kent State University for a review of the manuscript.

[4] Z. Sroubek, Z. Zdansky, and E. Simanek, Phys. Stat. Solidi 3, 165 (1963), observed in $BaTiO_3$ at 78°K a signal that they attribute to Pt^{3+}. Their crystals, however, were doped with about 0.4 at. % niobium, which acts like a donor, to reduce the valence state of diamagnetic Pt^{4+}. The fact that our samples contained only about 0.03 at. % niobium is additional evidence that the spectrum is not due to Pt^{3+}.

[5] In some crystals a weaker signal is observed, following the same spin-Hamiltonian as the strong one but with $g_{11}=2.00\pm0.01$ and $g_1=2.24\pm0.01$. This may be eventually attributed to Cu^{2+}.

[6] J. S. Griffith, *The Theory of Transition-Metal Ions* (Cambridge University Press, Cambridge, 1961), Chap. 12.

[7] J. E. Bennett, J. F. Gibson, and D. J. E. Ingram, Proc. Roy. Soc. (London) A240, 67 (1957).

[8] W. J. Nicholson and G. Burns, Phys. Rev. 129, 2490 (1963).

[9] F. Jona and G. Shirane, *Ferroelectric Crystals* (The Macmillan Company, New York, 1962), Chap. 5.

[10] E. S. Kirkpatrick, K. A. Müller, and R. S. Rubins (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Stimulated Phonon Emission by Supersonic Electrons and Collective Phonon Propagations

E. W. PROHOFSKY
Sperry Rand Research Center, Sudbury, Massachusetts

(Received 23 December 1963; revised manuscript received 27 January 1964)

The emission of phonons by electrons with supersonic drift velocities is examined, and the crystal momentum transferred to a piezoelectric phonon band is determined. This momentum is transferred principally to a band of phonons which are in a cone centered about the direction of electron drift and of frequencies approximately 10^{10}–10^{11} cps. This momentum-transfer process is stimulated in the sense that it amplifies any variations in the phonon momentum distribution. In particular, it will amplify a collective wave in the "phonon gas" which corresponds to a sound wave in a particle gas. The collective wave is described in terms of a local particle drift velocity and a local temperature. In a phonon gas, temperature plays the role that pressure plays in a particle gas. By the use of energy and momentum conservation equations, it is shown that such a collective wave is possible when the usual large losses are overcome. The derived gain and expected loss of the collective wave are shown to be comparable in CdS at room temperature. The current saturation observed when the electron drift velocity exceeds the sound velocity appears in the self-consistent solutions of the electron and phonon systems.

I. INTRODUCTION

THE observation of an anomalously slow phonon propagation has recently been reported in CdS crystals under conditions of acoustic gain.[1] This has been interpreted as a collective phonon propagation similar to the collective phonon wave which describes second-sound propagation in super-fluid helium. In nonconducting solids the collective phonon wave in a "phonon gas" could be considered analogous to a sound wave in a molecular gas. The collective wave would then be a density fluctuation in the phonon distribution, much like a sound wave is a particle density fluctuation in a particle gas. In a macroscopic description an increase in phonons gives rise to an increase in temperature. Temperature plays a role in phonon transport similar to pressure in particle gas transport. The collective phonon wave could be described equally well as a temperature fluctuation wave.

The propagation of such a collective phonon wave in a solid has been examined in some detail.[2] It was found that umklapp collisions (which have no counterpart in molecular gases) would destroy the crystal momentum of the phonons and damp the collective propagation. In solids, at all but lowest temperatures, this damping term is much larger than the harmonic term, and diffusion, (i.e., normal thermal conductivity) rather than harmonic propagation, occurs. At very low temperatures the umklapp collision rate is greatly reduced and collective waves may be possible, but the reduction in normal collisions would then tend to disrupt the waves. Normal collisions are necessary to maintain the collective nature of the wave, just as particle collisions are necessary for sound propagation in particle gases. These two conditions, long umklapp relaxation time and short normal relaxation time, are extremely difficult to satisfy simultaneously.

In a semiconducting solid, particularly those with piezoelectric phonon modes, one can transfer energy and crystal momentum from electrons to the phonons. It has been shown by Hutson, McFee, and White[3] that for ultrasonic waves, i.e., low-frequency phonons, the internal losses of the phonons can be more than compensated for when $v_d > v_s$. Here, v_d is the electron drift velocity and v_s the sound velocity. If such a mechanism could transfer crystal momentum to phonons capable of participating in a collective wave, which would overcome the internal losses, it would be possible to propagate the type of collective phonon wave discussed. Whether or not a particular phonon mode could participate in a collective wave would depend on how strongly coupled the mode was to the bulk of the collective wave by the mixing collisions.

The transfer of momentum or energy to phonons has been examined by White[4] for low frequencies, i.e., $q\Lambda < 1$, where q is the phonon wave vector and Λ is the electron free path. In this range, the electrons are able to equilibrate to their local surroundings. The best description of this system is in terms of the classical fields and currents used by White. For higher frequency phonons, where $q\Lambda > 1$, a better description of the system involves quantum mechanical electron and phonon states which undergo collisions. Pippard[5] has discussed the interaction mechanism for high-frequency phonons from a degenerate electron band with imposed drift velocity, $v_d > v_s$. Spector[6] has derived the amplification of high- and low-frequency phonons using a free-electron gas model for conduction of electrons in metals. Neither of these approaches is ideally suited to examine the highly nondegenerate-electron distribution which exists

[1] H. Kroger, E. W. Prohofsky, and R. W. Damon, Phys. Rev. Letters 11, 246 (1963).
[2] E. W. Prohofsky and J. A. Krumhansl, Phys. Rev. 133, A1403 (1964). E. W. Prohofsky, thesis, Cornell University, 1963 (unpublished).
[3] A. R. Hutson, J. H. McFee, and D. L. White, Phys. Rev. Letters 7, 237 (1961).
[4] D. L. White, J. Appl. Phys. 33, 2547 (1962).
[5] A. B. Pippard, Phil. Mag. 8, 161 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in CdS under conditions of acoustic gain. Paranjape[7] has discussed the amplification of high-frequency sound waves using a quantum mechanical approach suitable to nondegenerate electron distributions.

This paper will examine the transfer of momentum to high-frequency phonons using a golden-rule scattering integral technique suitable to nondegenerate-electron distributions in direct semiconductors. This momentum-transfer mechanism will, in turn, be used to evaluate the momentum transferred to collective phonon disturbances. It is believed that the high-frequency phonons will be most important in the collective wave behavior for two reasons: The number of higher frequency modes is much greater than that of low-frequency modes, and the mixing collision rate should be much less for low frequencies. This drop-off of mixing collision rate will supply a lower limit to the phonon frequencies involved in the collective wave. Those modes which have a lifetime longer than the period of the collective wave cannot be equilibrated within the collective wave.

In Sec. II, the first-order stimulated phonon emission and momentum-transfer rates are evaluated for a model in which there is a piezoelectric phonon mode and a nearly free, nondegenerate-electron band.

The results of Sec. II show that the electrons are strongly coupled to phonons of moderately high frequencies, but the interaction drops very rapidly beyond a certain cutoff frequency due to the mass mismatch between electrons and high-energy phonons. Energy and momentum transfer, upon collision, is most efficient for particles of roughly the same mass. An apparent phonon mass is

$$m_p = h\nu/v_s^2, \tag{1.1}$$

where h is Planck's constant, ν the frequency, and v_s the velocity of sound. The frequency most strongly coupled to electrons is that for which $m_p = m_e$, where m_e is the electron mass. This frequency is roughly $\nu = 10^{11}$ cps. The actual peak frequency found in Sec. II is very close to this value, but varies with temperature as it is effected by such factors as the available phonon and electron phase space and the average electron energy. The frequency dependence of the efficiency of energy and momentum transfer is greatly modified when the coupling has a frequency dependence, as is the case in deformation potential scattering.

The cutoff of the interaction for high-phonon frequencies enhances the probability of collective wave propagation. The phonon momentum losses rise sharply with frequency. Pumping momentum into a low-frequency band can give rise to a situation similar to that encountered in anomalous thermoelectric power.[8] The situation is illustrated in Fig. 1. A low-frequency band of phonons becomes "hot" with respect to the bulk of the thermal phonons. The losses from this band are slow because of a bottleneck which keeps these phonons from equilibrating with higher frequencies. The high-frequency umklapp relaxation time is very fast, and the thermal phonons remain uneffected in the propagation of the lower-frequency band.

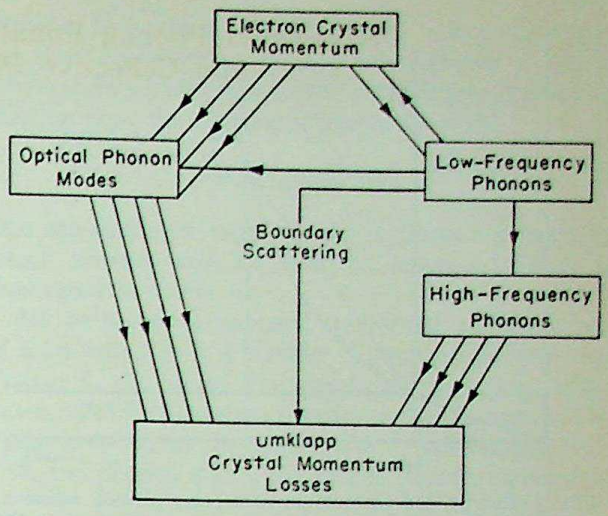


FIG. 1. Crystal momentum relaxation flow chart. Electrons accelerated by an electric field lose crystal momentum in collisions with optical and low-frequency acoustic phonons. For low electron drift velocities the relaxation to optical modes dominates. When $v_d > v_s$ the coupling to low-frequency acoustic modes becomes dominant. This low-frequency band can become hot since the momentum flowing into it can be greater than its losses to the high-frequency modes.

The emission process for $v_d > v_s$ is somewhat similar to Čerenkov radiation. As described in Sec. II, it, too, is limited to a forward cone centered about the direction of electron drift. The active phonon band which may support collective oscillations is a medium-frequency band of phonons occupying a cone in phonon-phase space centered about v_d.

The expression for the crystal-momentum transfer, found in Sec. II, is only valid in a region where $v_d \approx v_s$. The use of perturbation theory appears valid, as the dimensionless coupling parameter described by Froelich[9] is

$$\alpha = \left[\frac{\hbar}{2\rho\nu}\right]^{1/2}\left(\frac{ep}{d}\right)[\hbar\nu]^{-1} = \left[\frac{ep}{d}\right]\left[\frac{\pi}{\rho\hbar}\right]^{1/2}\nu^{-3/2}, \tag{1.2}$$

where ρ is the lattice density, e the electronic charge, p the proper piezoelectric constant, and d a dielectric constant. The parameter is certainly small for frequencies of the order of 10^{10}–10^{11} cps. For low values of electric field, the electron relaxation is mainly due to collisions with optical modes. This relaxation time is roughly 10^{-11}–10^{-12} sec, and the inequality $q\Lambda > 1$ is also satisfied for phonon frequencies of the order of 10^{11} cps.

The problem of the normal collision rate is discussed in Sec. III. The collisions between phonon states in the active band are primarily due to higher-order electron

[6] H. N. Spector, Phys. Rev. 127, 1084 (1962).
[7] B. V. Paranjape, Phys. Letters 5, 32 (1963).
[8] C. Herring, Phys. Rev. 96, 1163 (1954).
[9] H. Froehlich, Advan. Phys. 3, 325 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

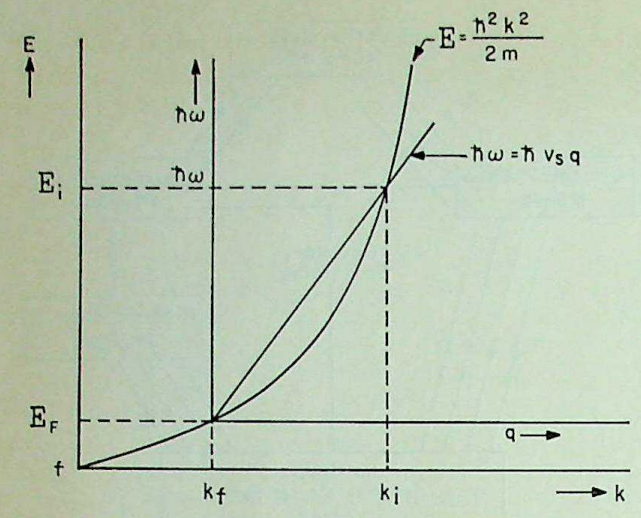


FIG. 2. Energy and momentum conserving phonon emission. The phonon dispersion relation $\hbar\omega=\hbar v_s q$ is superimposed on the electron dispersion relation $E=\hbar^2k^2/2m$. A possible phonon (ω,q) can be emitted in an energy and momentum conserving process if (1) $E_i=\hbar\omega+E_F$, and (2) $k_i=q+k_F$, where i and F refer to the initial and final electron states. All such double intersections insure that these two conditions are satisfied. The double intersections cease when the line is tangent to the curve when $E_i=E_F$ and the velocity of this state is equal to the vleocity of sound.

phonon collisions. The electrons, in this case, play a role very similar to that of the rotons in liquid helium. The same conditions which bring about large momentum transfer ($v_d>v_s$) cause this second-order collision mechanism to have a focusing effect on the phonons. That is, it tends to make the final phonon state more likely to be within the gain cone than the initial phonon state. The inability of the electron to transfer large amounts of energy to the phonons tends to keep the final state within the frequency limit of the active phonon band.

In Sec. IV, the equations for energy and crystal momentum conservation are applied to the active phonon band; the equation of conservation of momentum is applied to the electrons; and the interaction developed in Sec. II is used to couple the equations. A zero-order solution exists which describes the equilibrium distortion in the electron distribution (current), and the phonon distribution (heat flow) under the high electric field conditions. This solution shows a marked current saturation similar to that first reported by Smith.[10] In addition, solutions for displacements about this equilibrium exist which correspond to the collective propagation discussed. Throughout this section the excess energy stored in the electron distribution is ignored compared to that stored in the phonon distribution. Plasma effects are ignored because the electron density is assumed very small.

In Sec. V, numbers such as electron mass, piezoelectric constant, etc., are substituted into the equations that have been developed. The gain of the system is then compared to the expected phonon crystal momentum losses.

[10] R. W. Smith, Phys. Rev. Letters 9, 87 (1962).

II. STIMULATED PHONON EMISSION BY ELECTRONS IN PIEZOELECTRIC MATERIALS

In the quantum limit, i.e., $\mathbf{q}\cdot\mathbf{A}>1$, interactions involving electrons and phonons must conserve energy and crystal momentum modulo $\hbar\mathbf{G}$ ($\mathbf{G}$ being a reciprocal lattice vector). The possible initial and final states in which an electron emits or absorbs a single phonon in a normal collision (i.e., $\hbar\mathbf{G}=0$) can be determined by using a diagram similar to the one in Fig. 2. The curved line represents a dispersion relation for an electron in one dimension and is assumed to be a quadratic function of the momentum. The curve in Fig. 2 could be related to excited electrons in a direct material such as CdS. The straight line represents the acoustic phonon-dispersion relation; the slope equals the velocity of sound.

If the origin of the phonon line were placed at any point on the electron curve, a second intersection by the phonon line would imply conservation of energy and momentum between the initial and final electron states determined by the two intersections. All possible interactions correspond to some possible placements of the curves that result in two intersections. It can be seen that such intersections cease when the phonon line is just tangent to the electron curve. At this point, the energy and momentum of the emitted phonon have decreased to zero. The velocity of this limiting electron state is equal to the velocity of sound. For emission, the initial electron state must be on the high-energy side of the tangent point (i.e., $v_e>v_s$), and the final state must be on the low-energy side (i.e., $v_e<v_s$). For absorption, the initial and final states are reversed. In the case of a quadratic-electron-dispersion relation, the initial and final $\mathbf{k}$ values are equidistant from $\mathbf{k}'$, the location of the tangent point.

In a three-dimensional solid, the dispersion relation for electrons is a bowl-shaped hypersurface in four-

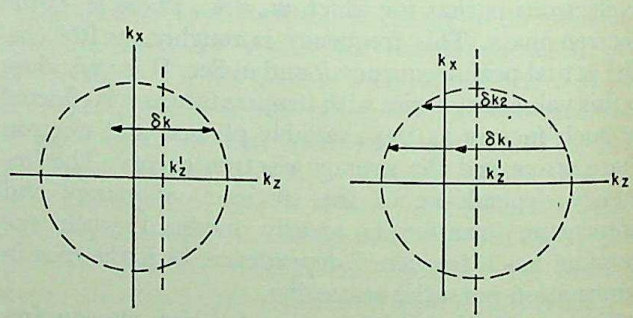


FIG. 3. Electron population inversion with respect to phonon emission. (This mechanism has been discussed by Pippard, Ref. 5.) The dashed circle represents phenomenologically the electron distribution in k space. The dashed line at k_z' represents those states whose velocity is equal to the sound velocity in the z direction. Emission or absorption of a phonon with $q=-\delta k$ involves initial and final states which are symmetrically spaced about the plane k_z'. In the figure on the left the electron distribution is at rest and there are fewer electrons on the high-energy side of k_z'. In the figure on the right the electron drift velocity is equal to the sound velocity and there are as many electrons on the high-energy side of k_z'. For greater drift velocities there will be more electrons on the high-energy side giving rise to a population inversion with respect to emission of phonons in the z direction.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

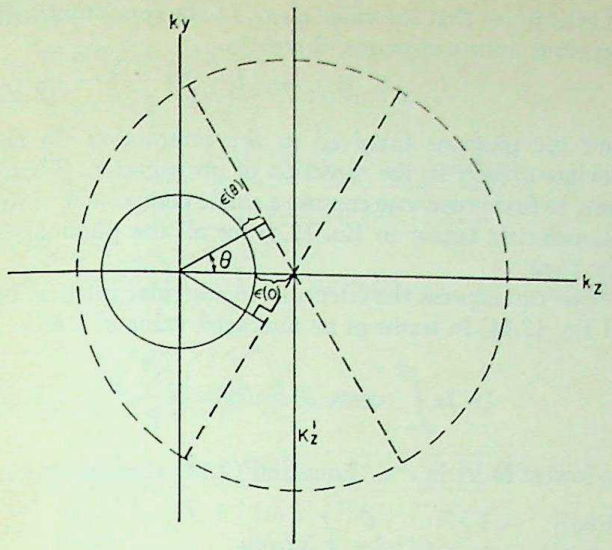


FIG. 4. The extent of population inversion for phonons not parallel to v_d. The solid-line sphere corresponds to the velocity of sound in any direction. The dashed sphere represents the electron distribution with a large drift velocity $v_d = v_s + \epsilon_0$. Phonons oriented at an angle θ will see a population inversion for $\epsilon(\theta) > 0$, where $\epsilon(\theta) = v_d \cos\theta - v_s(\theta)$.

dimensional space. Since only one phonon is involved in these processes, we may still use a one-dimensional phonon-dispersion line to examine the emission of phonons with one orientation. The origin of the phonon line is placed at a point on the electron-dispersion surface, and a second intersection implies a possible momentum and energy conserving process. The line and hypersurface will have a tangent point which will play a role similar to the tangent point in the one-dimensional analysis.

Based on the above discussion, it can be seen that a surface could be drawn in electron $\mathbf{k}$ space perpendicular to the direction of the emitted or absorbed phonon wave vector $\mathbf{q}$ at $\mathbf{k}'$, where $\mathbf{v}_e(\mathbf{k}') = \mathbf{v}_s$. All emission processes have the initial electron state on the high-energy side of this surface and the final state on the low-energy side. The change in electron wave vector, $\Delta\mathbf{k} = -\mathbf{q}$, is exactly bisected by this surface when the electron dispersion relation is quadratic in the momentum.

Pippard[5] has pointed out, in arguments based on a diagram similar to that of Fig. 3, that the phonon-emission rate could be greatly enhanced if the electron-drift velocity v_d were made greater than the sound velocity. In such a case, the centroid of the electron distribution in $\mathbf{k}$ space is outside the surface described in the preceding paragraph. Therefore, more electrons in states of high energy are opposite holes on the other side of this surface, and electron transition to the lower states is very probable. One can speak of a population inversion in $\mathbf{k}$ space with respect to phonon emission.

The emission of phonons for a given orientation in space will be shown as a first approximation to be proportional to a parameter ϵ, where

$$\epsilon = v_d - v_s, \tag{2.1}$$

where v_d is the electron drift velocity and v_s the sound velocity. For different phonon orientations, as those whose wave vector make an angle θ with the direction of electron flow, the emission is proportional to

$$\epsilon(\theta) = v_d \cos\theta - v_s. \tag{2.2}$$

One can define an angular cone in phonon q space in which phonon emission will be enhanced. This is demonstrated in Fig. 4.

All the collisions discussed are normal, and the effect of such collisions is a transfer of momentum from one system to the other. The population inversion, when $v_d > v_s$, favors momentum transfer to the phonons in the "gain cone" from the electrons. For well-defined electron and phonon states, the rate of phonon emission is

$$\left.\frac{\partial N_q}{\partial t}\right|_{\text{col}} = \frac{2\pi}{\hbar}\sum_{kk'} g^2\{[N(q)+1]f(k)[1-f(k')] - N(q)f(k')[1-f(k)]\} \times \delta(\mathbf{k}'-\mathbf{q}-\mathbf{k})\delta(\mathcal{E}_{k'}-\hbar\omega-\mathcal{E}_k), \tag{2.3}$$

where $N(q)$ is the occupation number of phonon mode q, $f(k)$ is the occupation number of the initial electron state, and $f(k')$ is the occupation number of the final electron state. The factor g is the coupling constant, or matrix element, for the coupling of the electron transition and the phonon. This energy can be written as

$$g = \frac{xep}{d} = \left(\frac{\hbar}{2\rho\omega_q}\right)^{1/2}\frac{ep}{d}, \tag{2.4}$$

where x is a normalized phonon displacement, e the charge on the electron, p the appropriate component of the piezoelectric tensor parallel to the phonon displacement, d the appropriate dielectric constant, ρ the lattice density, and ω_q the frequency of the phonon mode $\mathbf{q}$.

For high temperatures we assume $N(q) > 1$. This is equivalent to ignoring the spontaneous emission in comparison to the stimulated emission.

$$\left.\frac{\partial N}{\partial t}\right|_{\text{col}} \approx \frac{2\pi}{\hbar}\sum_{kk'} g^2 N[f(k)-f(k')] \times \delta(\mathbf{k}'-\mathbf{q}-\mathbf{k})\delta(\mathcal{E}_{k'}-\hbar\omega-\mathcal{E}_k). \tag{2.5}$$

For a quadratic electron dispersion relation,

$$k = (m^*/\hbar)v_s + (q/2) + k_{x,y}, \tag{2.6}$$

$$k' = (m^*/\hbar)v_s - (q/2) + k_{x,y}. \tag{2.7}$$

For drift velocities near the velocity of sound let

$$v_d = v_s + \epsilon. \tag{2.8}$$

One may assume that the electron distribution under high field is a Fermi distribution centered at $k_z = (m^*/\hbar)v_d$. That is,

$$f[(m^*/\hbar)v_d] = f_0(0), \tag{2.9}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where $f_0(k)$ is the equilibrium Fermi distribution. Substituting Eq. (2.8) into Eq. (2.5), and summing over k_x', k_y', k_z' to eliminate the $\delta(\mathbf{k})$ function,

$$\left.\frac{\partial N(q)}{\partial t}\right|_{\text{col}} = \frac{2\cdot 2\pi}{8\pi^3\hbar}\sum_k g^2 N(q) \times\left[f_0\left(\frac{q}{2}-\frac{m^*}{\hbar}\epsilon+k_{x,y}\right)-f_0\left(\frac{q}{2}+\frac{m^*}{\hbar}\epsilon+k_{x,y}\right)\right] \times\delta(\mathcal{E}_{k'}-\hbar\omega-\mathcal{E}_k)\approx\frac{-1}{2\pi^2\hbar}\sum_k g^2 N(q) \times\frac{\partial f_0}{\partial E}\,\text{grad}_k E\cdot\frac{m^*}{\hbar}\cdot 2\epsilon\delta(\mathcal{E}_{k'}-\hbar\omega-\mathcal{E}_k) \quad (2.10)$$

as f_0 has reflection symmetry.

Summing over k_z to eliminate $\delta(\mathcal{E})$,

$$\left.\frac{\partial N(q)}{\partial t}\right|_{\text{col}} = \frac{1}{2\pi^2\hbar}\int_{k_{x,y}} g^2 N(q) f_0\left(\frac{q}{2}+k_{x,y}\right) \times\left[1-f_0\left(\frac{q}{2}+k_{x,y}\right)\right]\frac{1}{KT}\frac{m^*}{\hbar}2\epsilon dk_x dk_y. \quad (2.11)$$

After replacing the remaining sums by integrals and carrying out the integration over $dk_x dk_y$ we obtain

$$\left.\frac{\partial N(q)}{\partial t}\right|_{\text{col}} = \frac{1}{2\pi}\frac{1}{\hbar^4}m^{*2}g^2 N(q)\epsilon \times\left[1-\tanh\left(\frac{\hbar^2q^2}{16m^*KT}-\frac{E_F}{2KT}\right)\right]. \quad (2.12)$$

The total crystal momentum transferred via this process is

$$\left.\frac{\partial}{\partial t}\langle\hbar\mathbf{q}\rangle\right|_{\text{col}} = \sum_q \hbar\mathbf{q}\left|\frac{\partial N(q)}{\partial t}\right|_{\text{col}}. \quad (2.13)$$

Replacing the summation by an integral, Eq. (2.13) becomes

$$\left.\frac{\partial\langle\hbar\mathbf{q}\rangle}{\partial t}\right|_{\text{col}} = \frac{1}{16\pi^4\hbar^4}\int m^2g^2\hbar q\cos\theta N(q)\epsilon(\theta) \times\left[1-\tanh\left(\frac{\hbar^2q^2}{16m^*KT}-\frac{E_F}{2KT}\right)\right]2\pi q^2\sin\theta d\theta dq. \quad (2.14)$$

The interesting range of θ runs from zero to θ', where θ' describes a gain cone in $\mathbf{q}$ space, and

$$\epsilon(\theta') = (v_s+\epsilon)\cos\theta' - v_s = 0, \qquad \cos\theta' = \frac{1}{1+\epsilon/v_s}. \quad (2.15)$$

It is assumed that the value $\epsilon/v_s \ll 1$ to be consistent with previous approximations, therefore

$$\theta' \approx (2\epsilon/v_s)^{1/2} < 1, \quad (2.16)$$

and the phonons involved in the interaction do not deviate greatly in the direction of propagation. Therefore, to first order, one can use a single component of the piezoelectric tensor in Eq. (2.4) for all the phonons in the cone.

One can express the effects of the angular integration of Eq. (2.14) in terms of an averaged value of $\epsilon(\theta)$.

$$\bar{\epsilon} = 2\pi\int_0^{\theta'}\cos\theta\epsilon(\theta)\sin\theta d\theta \approx 2\pi\frac{\epsilon^2}{v_s}, \quad (2.17)$$

to lowest order in ϵ/v_s. Equation (2.14) then becomes

$$\left.\frac{\partial\langle\hbar q\rangle}{\partial t}\right|_{\text{col}} = \frac{1}{8\pi^2}\frac{1}{\hbar^4}m^{*2}g^2\frac{\epsilon^2}{v_s}\int N(q)\hbar q \times\left[1-\tanh\left(\frac{\hbar^2q^2}{16m^*KT}-\frac{E_F}{KT}\right)\right]q^2dq. \quad (2.18)$$

For a nondegenerate electron gas the term in the square brackets can be simplified as

$$1-\tanh\left(\frac{\hbar^2q^2}{16m^*KT}-\frac{E_F}{2KT}\right) \approx\frac{\hbar^3n}{2(2m^*KT)^{3/2}}\frac{1-\tanh(\hbar^2q^2/16m^*KT)}{1+\tanh(\hbar^2q^2/16m^*KT)}, \quad (2.19)$$

where n is the electron density. A further simplification occurs if we define a temperature T_s by

$$KT_s = \tfrac{1}{2}m^*v_s^2; \quad (2.20)$$

and Eq. (2.18) becomes

$$\left.\frac{\partial\langle\hbar q\rangle}{\partial t}\right|_{\text{col}} \approx\frac{\pi}{8}\frac{1}{\hbar}m^{*2}g^2v_s^3(KT_sKT)^{-3/2}\frac{\epsilon^2}{v_s}n\int\hbar qN(q) \times\frac{[1-\tanh(\hbar^2v_s^2q^2/32KT_sKT)]q^2dq}{1+\tanh(\hbar^2v_s^2q^2/32KT_sKT)}. \quad (2.21)$$

We now allow for a distortion in the phonon field characterized by the phonons within the gain cone having a net drift velocity $\boldsymbol{\lambda}$. In this case, the distorted phonon probability distribution becomes

$$N(q) = N_0(q)+\boldsymbol{\lambda}\cdot\hbar\mathbf{q}\frac{1}{KT}N_0(N_0+1), \quad (2.22)$$

where N_0 refers to the equilibrium Bose distribution. It is now possible to separate the momentum transfer process into two parts, i.e.,

$$\left.\frac{\partial\langle\hbar q\rangle}{\partial t}\right|_{\text{col}} = R_1\frac{\epsilon^2}{v_s}+R_2\frac{\epsilon^2\lambda}{v_s^2}, \quad (2.23)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$R_1=\frac{\pi^2}{2}\frac{m^{*2}}{\rho v_s}\left(\frac{ep}{d\hbar}\right)^2\left(\frac{T}{T_s}\right)^{3/2}n \times\int x^2\frac{1}{e^x-1}\frac{1-\tanh[(1/32)(T/T_s)x^2]dx}{1+\tanh[(1/32)(T/T_s)x^2]} =\frac{\pi^2}{2}\frac{m^{*2}}{\rho v_s}\left(\frac{ep}{d\hbar}\right)^2\left(\frac{T}{T_s}\right)^{3/2}n\int r_1(x)dx, \quad (2.24)$$

$$R_2=\frac{\pi^2}{8}\left(\frac{ep}{d\hbar}\right)^2\frac{m^{*2}}{\rho v_s}\left(\frac{T}{T_s}\right)^{3/2}n \times\int x^3\,\mathrm{csch}^2\frac{x}{2}\frac{1-\tanh[(1/32)(T/T_s)x^2]dx}{1+\tanh[(1/32)(T/T_s)x^2]} =\frac{\pi^2}{2}\left(\frac{ep}{d\hbar}\right)^2\frac{m^{*2}}{\rho v_s}\left(\frac{T}{T_s}\right)^{3/2}n\int r_2(x)dx, \quad (2.25)$$

where

$$x=\hbar v_s q/KT. \quad (2.26)$$

The first term on the right-hand side of Eq. (2.23) describes spontaneous emission—spontaneous in the sense that the drifting electron distribution loses momentum to a nondrifting phonon distribution. The second term is a stimulated-emission term—stimulated in the sense that momentum is transferred at a rate determined by the net momentum present in the phonon system. This stimulated term describes a process whose net effect is a linear amplification of the net local momentum of the phonon distribution. The function $r_2(x)$ is greater than $r_1(x)$, and the "stimulated" emission term can be comparable to the "spontaneous" one.

The function, $(T/T_s)^{3/2}r_2(\nu)$, is plotted in Fig. 5 for several values of T/T_s. The values of T/T_s correspond to helium, nitrogen, and room temperature for free electron masses, and room temperature for an effective mass of $0.2m_0$. At low-phonon frequencies $q\Lambda<1$ this quantum approach is not valid. A more valid approach at these low frequencies is White's,[4] as previously discussed. The increase in the gain for phonons (as T/T_s increases) reflects the increased stimulation due to the presence of greater numbers of phonons. The shift in the frequency of maximum gain reflects the fact that the electrons have more energy and are more capable of emitting higher energy phonons.

III. NORMAL PHONON COLLISIONS AND HIGHER ORDER ELECTRON PHONON COLLISIONS

Phonons undergo collisions with other phonons via the anharmonic terms of the interatomic forces. Collisions in which no net crystal momentum is lost are called normal collisions. Such normal collisions are necessary for the propagation of a collective wave and establishing a local temperature. The collective wave

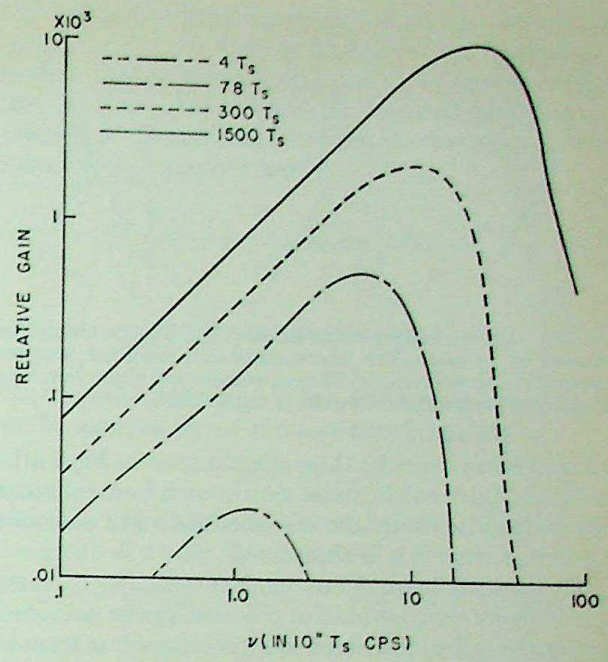


FIG. 5. Relative stimulated momentum gain as a function of phonon frequency. The function $r_2(\nu)$ is plotted for $T=4T_s$, $78T_s$, $300T_s$, and $1500T_s$. These temperatures correspond roughly to liquid helium, liquid nitrogen, and room temperature for $m^*\approx m_e$, and room temperature for $m^*\approx 0.2m_e$.

examined in this paper, however, must be propagated within a restricted band of lower-than-thermal frequencies, since the momentum loss of thermal phonons is much greater than the momentum gain available to the system. As pointed out in Sec. I, only the slow rate of equilibration of these low frequencies to the thermal reservoir makes the collective wave possible. Normal collisions in which momentum is transferred to phonon states outside of this restricted band will dissipate the momentum of the collective wave and give rise to damping. Instead of colliding among themselves, the low-frequency phonons within the restricted band are more likely to collide via the anharmonic perturbations with thermal-frequency phonons due to the much larger number of available states at higher frequencies. These collisions play the same role in this restricted-band collective wave as do the umklapp collisions in thermal phonon collective waves. The rate of these processes proceeding in the absence of carriers, i.e., dark attenuation in photoconductors, will be an estimate of the momentum losses from the restricted band to the thermal reservoir.

The slow rate of relaxation of the hot phonon band gives rise to a relaxation bottleneck. The phonons within this hot band are more likely to undergo further collisions with electrons before relaxation can occur. Successive electron-phonon collisions or higher order electron-phonon collisions are the fastest mixing collisions in the band and will set up a somewhat distorted equilibrium within the hot phonon band. The essential mixing collisions necessary for the propagation of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

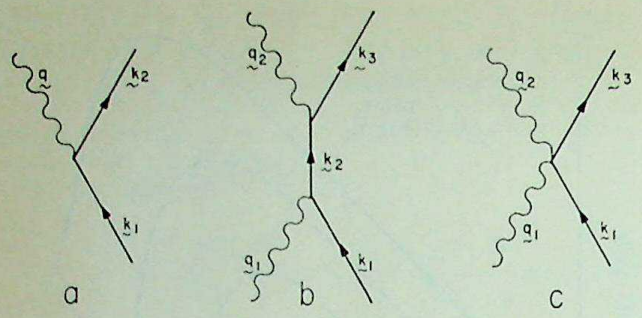


FIG. 6. (a) First-order phonon emission. (b) Phonon absorption followed by emission. The intermediate state need not conserve energy if it is short-lived. (c) Phonon emission and absorption. The existence of an intermediate state is not specified.

collective wave will be those diagrammed in Figs. 6(b) and 6(c). There will be some contribution from collisions [Fig. 6(b)] in which the intermediate state need not conserve energy if it is short lived.

The effect of large electronic drift velocities on these second-order electron-phonon processes can be examined qualitatively by describing the processes in a manner analogous to the emission problem of Sec. II. The change in electronic configuration in these second-order processes is the same as it would be if the process were an emission or absorption of an apparent phonon with energy and momentum equal to the differences in energy and momentum between the initial and final phonon states. The enhancement of a particular interaction occurs where $v_d > v_a$, where v_a is the apparent emitted phonon velocity and is

$$v_a = \frac{\Delta\omega}{\Delta q} = \frac{v_s}{1+(q_1/\Delta q)[1-\cos(\varphi-\theta)]}, \quad (3.1)$$

where q_1, Δq, φ, and θ are illustrated in Fig. 7. A gain cone extending over an angle φ' for these second-order processes can be defined where

$$\cos\varphi' = \cos\theta' \frac{1}{1+(q_1/\Delta q)[1-\cos(\varphi'-\theta)]}, \quad (3.2)$$

where $\cos\theta'$ is given in Eq. (2.15). Because of the electron population inversion, all interactions involving states q_2 whose vectors terminate on the ω_2 line inside φ' will tend to be apparent emissions, since the net interaction is more probable if q_2 is the final state and q_1 the initial state. For q_2 terminating outside φ' the net interaction favors transitions with q_2' as initial state and q_1 as final state. On the basis of this analysis it is seen that the mixing collisions tend to keep the excess phonon population within the gain cone where it is amplified, rather than outside the gain cone where it could be quickly damped. Furthermore, the focusing effect of the second-order collisions contribute to the net momentum gain of the phonon distribution.

One can easily obtain a lower limit of the second-order scattering from those intermediate states k_2 [Fig. 6(b)] in which energy is conserved. The second-order matrix element from second-order perturbation theory is

$$(k_3q_2|V^2|k_1q_1) = \sum_{k_2} \frac{(k_2|V^1|k_1q_1)(k_3q_2|V^1|k_2)}{E(k_1)+\hbar\omega_1-E(k_3)-i/2\Gamma}, \quad (3.3)$$

where Γ is an energy related to the finite lifetime of the intermediate state. That group of states k_2 in which energy is conserved has a resonant energy denominator, and the matrix element takes on a Breit-Wigner form. In the case of strongly enhanced first-order processes, the main part of the broadening will come from processes like "a" in Fig. 6. In this situation, the contribution to the scattering by the energy-conserving intermediate states will be approximately equal to the rate of first-order absorption processes. The final states of these higher order processes will lie in the same frequency band as the first-order emission processes. The contribution to mixing, from processes in which the intermediate state is not an energy conserving one, is expected to be small. This is because of the small size of the parameter α defined in Eq. (1.2) which determines the relative size of terms in this perturbation expansion.

A better understanding of the relative magnitudes of the emission and absorption rates can be achieved by examination of Fig. 3. When $v_d = v_s$ the electron distribution is symmetrically distributed such that the probability of absorption and emission are equal. The rate of both processes is very large because the energy and momentum in both the electron and phonon systems are ideally distributed for such interactions. The net transfer of momentum or energy is zero as the two processes cancel each other. An increase in drift velocity gives rise

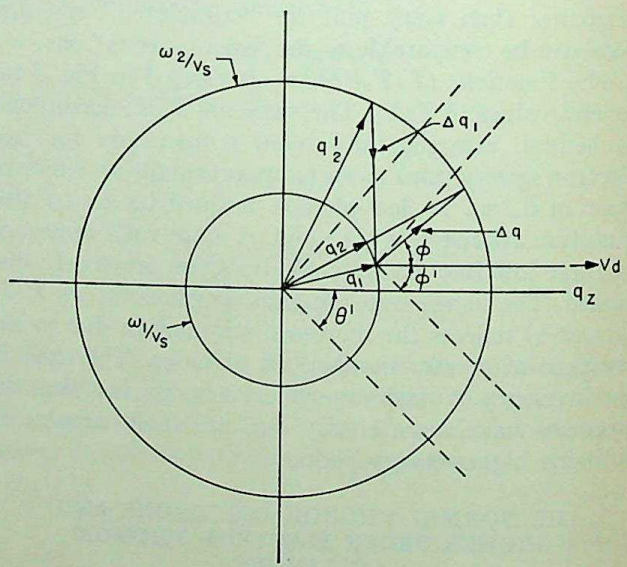


FIG. 7. Higher order electron phonon processes. These processes can be described in terms of the emission or absorption of apparent phonons. Two such processes are shown which have different changes in wave vector q_2-q_1 or $q_2'-q_1$ for a fixed energy difference $\omega_2-\omega_1$. This is equivalent to a variable apparent velocity. The cone φ' is the gain cone for these apparent phonons. Emission or absorption of these apparent phonons will be more probable depending on whether q_2 is inside or outside φ'. θ' is the gain cone for first-order processes.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to a net phonon gain by increasing the emission rate relative to the absorption rate. The mixing rate, which is limited by the slower of the two processes (absorption or emission), would then be decreased due to the decrease in the absorption rate. If successive first-order processes are the dominant contribution to the mixing collisions, an increase in net gain will be accompanied by a decrease in the mixing rate. The ability to propagate a collective wave will depend on achieving a delicate balance between the gain and the mixing rate.

Higher temperatures will broaden the electron distribution and therefore reduce the sharpness of the population inversion for a given electron drift velocity. The phonon momentum gain per phonon mode actually decreases with increasing temperature [see Eq. (4.35)]. Higher temperatures will therefore tend to increase the mixing collision rate; lower temperatures will increase the gain. For a given drift velocity there is a temperature in which the two conflicting requirements for collective propagations may be simultaneously satisfied. Higher temperature will reduce the gain, and lower temperatures will reduce the mixing rate.

IV. CURRENT SATURATION AND COLLECTIVE WAVES

An equation for conservation of electron momentum can be derived from the Boltzmann transport equation and written as

$$\frac{\partial}{\partial t}\langle \hbar k_i\rangle+\frac{\partial}{\partial x_j}\langle v_j \hbar k_i\rangle-\frac{Ee}{\hbar}\left\langle\frac{\partial f}{\partial k_i}\right\rangle=\Delta\hbar q\Big|_{\text{col}}, \quad (4.1)$$

where the brackets refer to expectation values over the electron distribution. For a spatially uniform electron distribution, the second term is zero. When the drift velocity is less than the velocity of sound, the collision term can be simplified by a relaxation time approximation and becomes

$$\Delta\hbar q|_{\text{col}}=(-nmv_d/\tau_0), \quad (4.2)$$

where nm is the total electron mass per unit volume and τ_0 is the electron relaxation time. For materials such as CdS at room temperature, the electron relaxation is mainly due to collision with optical phonon modes. That part of the relaxation which is to optical modes, should remain relatively unaffected when v_d exceeds v_s. If we add the collision terms of Eq. (2.23) to represent the greatly enhanced part of the relaxation, Eq. (4.1) becomes

$$\frac{\partial}{\partial t}nmv_d-\frac{Ee}{\hbar}\left\langle\frac{\partial f}{\partial k}\right\rangle=-\frac{nmv_d}{\tau_0}-R_1\frac{\epsilon^2}{v_s}-R_2\frac{\epsilon^2\lambda}{v_s^2}; \quad (4.3)$$

and, using the usual definition of mobility μ, this becomes

$$\frac{\partial}{\partial t}nmv_d-\frac{nm\mu E}{\tau_0}=-\frac{nmv_d}{\tau_0}-R_1\frac{\epsilon^2}{v_s}-R_2\frac{\epsilon^2\lambda}{v_s^2}. \quad (4.4)$$

Equations for the conservation of energy and crystal momentum for the "hot" phonon band can be similarly derived. The conservation of energy is needed in this case, as the phonons are the thermal and inertial reservoir of the combined electron-phonon system. The conservation equations are

$$\frac{\partial}{\partial t}\langle\hbar q_i\rangle+\frac{\partial}{\partial x_j}\langle v_j\hbar q_i\rangle=\langle\Delta\hbar q_i\rangle_{\text{col}}, \quad (4.5)$$

$$\frac{\partial}{\partial t}\langle\hbar v_s q\rangle+\frac{\partial}{\partial x_j}\langle v_j\hbar v_s q\rangle=\langle\Delta\hbar v_s q\rangle_{\text{col}}. \quad (4.6)$$

The variation of energy in the restricted phonon band can be specified by an effective temperature

$$T=T_0+T_1, \quad (4.7)$$

where T_1 may fluctuate. The distribution function for this group of modes then becomes

$$N(q)=[\exp|\hbar v_s q-\hbar\lambda\cdot\mathbf{q}/K(T_0+T_1)|-1]^{-1}, \quad (4.8)$$

where λ is the phonon drift described in Sec. II. Using Eq. (4.7) and (4.8), Eqs. (4.5) and (4.6) become

$$\frac{\partial}{\partial t}M\lambda+\frac{\partial}{\partial z}\gamma CT_1=R_1\frac{\epsilon^2}{v_s}+R_2\frac{\lambda\epsilon^2}{v_s^2}-\frac{M}{\tau_\lambda}\lambda, \quad (4.9)$$

$$\frac{\partial}{\partial t}CT_1+\frac{\partial}{\partial z}D\lambda_1=v_s(\Delta\hbar q)_{\text{col}}. \quad (4.10)$$

Here, τ_λ refers to collisions which transfer momentum out of the active phonon band, C is the specific heat of the phonon band, γ is the average of $\cos^2\theta$ over the cone of phonons making up the phonon band, M and D are constants which depend on the ambient temperature as

$$M=\frac{1}{8\pi^3}\int\hbar q\cos\theta\frac{\cos\theta\hbar q}{KT}N_0(q)[N_0(q)+1]q^2\times\sin\theta dq d\theta d\varphi$$

$$=\frac{1}{4\pi^2}\frac{\gamma(KT_0)^4}{\hbar^3 v_s^5}\int x^4\operatorname{csch}^2\frac{x}{2}dx\int_0^{\theta'}\sin\theta d\theta, \quad (4.11)$$

$$D=\frac{1}{4\pi^2}\frac{\gamma(KT_0)^4}{\hbar^3 v_s^3}\int x^4\operatorname{csch}^2\frac{x}{2}dx\int_0^{\theta'}\sin\theta d\theta. \quad (4.12)$$

It is understood that the integral over x includes only those frequencies in the "hot" phonon band.

A static equilibrium solution can be found by setting time and space derivatives equal to zero and solving for ϵ_0 and λ_0. Equations (4.4) and (4.9) become

$$\frac{-nm}{\tau_0}\mu E=\frac{-nm}{\tau_0}(v_s+\epsilon_0)-\frac{R_1\epsilon_0^2}{v_s}-\frac{R_2\epsilon_0^2\lambda_0}{v_s^2}, \quad (4.13)$$

$$0=\frac{R_1\epsilon_0^2}{v_s}+R_2\frac{\epsilon_0^2\lambda_0}{v_s^2}-\frac{M\lambda_0}{\tau_\lambda}. \quad (4.14)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The sum of these equations is

$$\frac{nm}{\tau_0}\beta=\frac{nm}{\tau_0}\frac{\epsilon_0}{v_s}+\frac{M}{\tau_\lambda}\frac{\lambda_0}{v_s}, \tag{4.15}$$

where

$$\beta=\frac{\mu E}{v_s}-1 \tag{4.16}$$

and is the effective excess field driving the system. The righ-hand side of Eq. (4.15) consists of the net loss from the combined electron-phonon system due to electrons and phonons colliding outside the system. The resulting cubic equation for ϵ_0 or λ_0 can be easily solved in two limits: for $0<\beta\ll 1$

$$\epsilon_0/v_s=\tfrac{1}{2}\beta, \tag{4.17}$$

and for $\beta\approx 1$

$$\epsilon_0/v_s=(R_2/2R_1)^{1/2}\beta^{1/2}. \tag{4.18}$$

The current saturation, observed in CdS crystals[10] when the electron-drift velocity is comparable to or exceeds the sound velocity, is explained by these equations. At low electric field ($\alpha<0$) the coupling between Eqs. (4.13) and (4.14) is negligible. The primary electron damping is due to the optical phonon modes and normal resistive behavior results. When $v_d\approx v_s$ the further current increases Δj for increases in field are given by

$$\Delta j=ne\Delta\epsilon_0=ne\tfrac{1}{2}\mu\Delta E. \tag{4.19}$$

This causes a knee in the current voltage curves where the slope drops to one-half its original value. At very high electric fields the saturation increases and Δj becomes proportional to the square root of E. Accompanying this saturation, a net drift is set up in the phonon distribution described by λ_0, where

$$\frac{\lambda_0}{v_s}=\frac{nm\tau_\lambda}{M\tau_0}\left(\beta-\frac{\epsilon_0}{v_s}\right). \tag{4.20}$$

Although the electrons are capable of conforming to the instantaneous phonon distribution very quickly, it will take a time, roughly equal to the phonon relaxation time, for the phonons to achieve their equilibrium value. Such a delay in achieving equilibrium can explain the observed transient behavior in CdS.[10]

It is emphasized that these solutions are valid only in bulk material. No attempt has been made to introduce the effect of boundary conditions or spatial variations. The size of the sample should certainly effect transient time, particularly for small crystals. Also, no attempt has been made to include the effects of the low-frequency $q\Lambda<1$ phonons on saturation. It is felt, however, because of the greater number of available high-frequency modes, that the effect of the low-frequency phonons will not be too important.

The electron-phonon system can be examined for its behavior under slight departures from equilibrium. The parameters ϵ, λ, and T can be allowed to vary in both time and space about their equilibrium values, i.e., $\epsilon=\epsilon_0+\epsilon_1$, $\lambda=\lambda_0+\lambda_1$, $T=T_0+T_1$, where

$$\begin{aligned}\epsilon_1&=\epsilon_1{}^0e^{i(\Omega t-lz)},\\ \lambda_1&=\lambda_1{}^0e^{i(\Omega t-lz)},\\ T_1&=T_1{}^0e^{i(\Omega t-lz)}.\end{aligned} \tag{4.21}$$

Equations (4.4, 4.9, and 4.10) become

$$i\Omega nm\epsilon_1{}^0=-A\lambda_1{}^0-B\epsilon_1{}^0-(nm/\tau_0)\epsilon_1{}^0, \tag{4.22}$$

$$i\Omega M\lambda_1{}^0-il\gamma CT_1{}^0=B\epsilon_1{}^0+A\lambda_1{}^0-(M/\tau_\lambda)\lambda_1{}^0, \tag{4.23}$$

$$i\Omega CT_1{}^0-ilD\lambda_1{}^0=v_s[B\epsilon_1{}^0+A\lambda_1{}^0-(M/\tau_\lambda)\lambda_1{}^0], \tag{4.24}$$

where

$$A=R_2(\epsilon_0{}^2/v_s{}^2) \tag{4.25}$$

$$B=2[R_1(\epsilon_0/v_s)+R_2(\epsilon_0\lambda_0/v_s{}^2)]. \tag{4.26}$$

The dispersion relation for simultaneous solution of Eqs. (4.22)–(4.24) has roots which describe the collective propagation discussed. This dispersion relation is

$$-\Omega^2-2i\Omega[A'-(1/\tau_u)]+\gamma\bar{v}_s{}^2l^2=0, \tag{4.27}$$

where

$$A'=A/M, \tag{4.28}$$

$$B'=B/nm, \tag{4.29}$$

and

$$\bar{v}_s{}^2=D/M. \tag{4.30}$$

Substituting in Eqs. (4.25), (4.11), and (2.26), this becomes

$$A'=8\pi^4\left(\frac{ep}{d}\right)^2\frac{\hbar v_s{}^4m^{*2}n}{(KT)^4\rho}\left(\frac{T}{T_s}\right)^{3/2}n \times\frac{\epsilon_0}{v_s}\frac{\int x^3\,\mathrm{csch}^2\frac{x}{2}\,\frac{[1-\tanh(1/32)(T/T_s)x^2]dx}{1+\tanh[(1/32)(T/T_s)x^2]}}{\int x^4\,\mathrm{csch}^2\frac{x}{2}dx}, \tag{4.31}$$

where the approximation

$$\int\cos^2\theta\sin\theta d\theta\approx\epsilon/v_s \tag{4.32}$$

has been used. Equation (4.32) holds for small ϵ/v_s. A simplifying approximation can be made in Eq. (5.1) by assuming

$$[1-\tanh y]=\begin{cases}1 & y<0.5\\ 0, & y>0.5\end{cases}, \tag{4.33}$$

and expanding the $\mathrm{csch}x/2$ functions for small values of x. The maximum value of x in the integrals is then given by

$$x_{\max}=4(T_s/T)^{1/2}. \tag{4.34}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Carrying out the integral to x_{max}, Eq. (4.31) becomes

$$A' = 3\pi^4 \left(\frac{ep}{d}\right)^2 \frac{\hbar v_s^4 m^{*2}}{(KT)^4 \rho} \left(\frac{T}{T_s}\right)^2 n \frac{\epsilon_0}{v_s}. \quad (4.35)$$

The damping of the collective wave is

$$\omega_I = \tfrac{1}{2}(1/\tau_\lambda - A'), \quad (4.36)$$

where A is the stimulated momentum gain rate for a phonon distortion due to phonon emission by the electrons. When $\omega_I < 0$ a phonon distortion will grow rather than be damped.

In a more detailed treatment of this problem, one would retain the q dependence of both A' and $1/\tau_\lambda$. The gain band would be defined as those phonon modes for which

$$(1/\tau_\lambda)(\mathbf{q}) - A'(\mathbf{q}) < 0. \quad (4.37)$$

The loss due to $1/\tau_\lambda(\mathbf{q})$ will modify the frequency distribution of the gain band. The value of gain at various frequencies shown in Fig. 5 is greater than the net gain by an amount equal to $1/\tau_\lambda(\omega)$, since the losses increase with increasing frequency and the density of frequencies participating in collective waves will be shifted to lower frequencies. The angular distribution of the hot phonons will also be affected since the gain cone should be defined by that angle at which the net gain is zero.

An exact determination of the collective wave velocity would require a more thorough investigation of the role of the mixing collisions as well as detailed knowledge of the loss mechanisms which damp the wave. The problem is in knowing the extent of the cone over which $\cos^2\theta$ must be averaged in the determination of γ of Eq. (4.9). The mixing collisions bring about transfers of momentum which involve phonon modes well outside the gain cone. The effect of the drag of this additional specific heat on the modes involved in the transport of momentum is not as yet understood. If the entire forward half of the phonon distribution is involved in the local thermalization process [i.e., determining T_1 of Eq. (4.9)], the value of γ will be $\frac{1}{3}$. A restriction to some smaller cone would increase this value. The maximum value of γ is $\gamma = 1$, and this would occur when the cone is limited to a line in the direction of current flow. The absence of mixing collisions would also give rise to an apparent value of $\gamma = 1$.

The factor $\bar{v}_s^2$ of Eq. (4.9) should be a weighted average over all the excitations participating in the collective wave. This is important in the case of second-sound propagation in He II as the admixture of rotons slows the collective wave velocity.[11] In the wave discussed in this paper, the electrons participate in the collective wave and should be averaged in. The effect of the electrons should be very small due to their relatively small numbers and small mass. The phonons involved in the wave also have varying velocities. This is because the phonons having piezoelectric fields become impure modes in some directions. In CdS the piezoelectric mode along the "C" axis is a longitudinal mode. The effect of the different velocities of phonons for large angles, off the direction of propagation, is reduced by the factors of $\cos\theta$ in the constants M and D of Eqs. (4.11) and (4.12). This change in velocity with orientation would also effect the extent of the gain cone, etc., with orientation.

[11] K. R. Atkins, *Liquid Helium* (Cambridge University Press, Cambridge, England, 1959), p. 148.

V. CONCLUSION

The observation of an anomalously slow acoustic propagation in CdS has been recently reported.[1] The effect is directly related to the passage of a piezoelectrically active shear wave through the crystals under conditions of acoustic gain. The anomaly could only be explained as a wave traveling in the CdS with a velocity just slightly faster than $1/\sqrt{3}v_s$, where v_s is the velocity of the piezoelectric shear wave. This shear wave is the slowest of the possible sonic waves in CdS. The pulses were observable only with applied voltages such that the electrons had drift velocities roughly 50 percent greater than the velocity of sound. For drift velocities, roughly twice the velocity of sound, the pulses were damped out. Similarly, the pulses were observed only over a small range in temperature near 300°K. Higher and lower temperatures caused the wave to be damped. At low temperatures the velocity of the wave was observed to increase as the wave was damped. These observations can be explained on the basis of the theoretical development of a collective-wave propagation presented above.

The collective wave consists of a harmonic fluctuation in the density of phonons within an active phonon band (phonon frequencies $\sim 10^{10}$–10^{11} sec^{-1}). This could also be described as a harmonic temperature fluctuation in the active band. The phonon density (temperature) fluctuations in turn, have a frequency equal to that of the shear wave which generated them (shear wave frequency $\sim 10^7$ sec^{-1}). The role of the initial shear wave was to introduce a single-frequency perturbation, which when amplified, could be coherently detected.

The propagation of a collective wave requires that two conditions be satisfied: (1) The gain in crystal momentum of the active phonon band must exceed the losses, and (2) mixing collisions among phonons within the band must be sufficiently frequent that a local temperature can be defined within a collective wavelength.

In Sec. II of this paper, the momentum gain for piezoelectrically active phonons has been evaluated. The gain mechanism is such that the electrons couple only to a band of phonons of frequencies 10^{10}–10^{11} cps. In CdS, under conditions approximating those of the experimental observation, (i.e., at room temperature with $n = 10^{13}$ and $p \approx 10^7$ D/esu)[12] the gain in the band given by Eq. (4.35) is

$$A' \approx 10^3(\epsilon/v_s)\ \text{sec}^{-1}. \quad (5.1)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

That is, this gain is great enough, at large electric field strengths, to overcome active phonon-thermal relaxation times of 10^{-8} sec.

The losses which this gain would have to overcome to satisfy condition (1) are the losses from the active phonon band to thermal-phonon reservoir. These processes are due to the anharmonic perturbations and do not involve electrons. The phonon-thermal phonon relaxation time for the phonon frequencies expected to be of importance is difficult to estimate, and direct measurements at these frequencies and temperatures in CdS are not available. The difficulty arises because the active band lies between very high ultrasonic frequencies and very low thermal frequencies. Woodruff and Ehrenreich[13] have shown that the phonon-phonon damping increases roughly as ω^2 in the ultrasonic region, and as ω for higher frequencies. The extrapolation to the frequencies important in collective wave propagation could be approximated by a variation as ω^n with $1<n<2$.

The dark relaxation time (i.e., no carriers) measured by Kroger[14] in CdS crystals, which have shown collective wave propagation, is 10^{-3}–10^{-4} sec^{-1} at 10^7 cps. If one extrapolates from 10^7 cps to the higher frequencies of the active phonon band, a very large part of the phonon band which is capable of supporting collective waves [satisfying condition (2)] will have stimulated gain greater than the losses. The net gain would be the difference between two large numbers, roughly the order of 10^8 sec^{-1}. From Eq. (4.35) it can be seen that the momentum gain decreases for decreasing electric field and for higher temperatures. This decrease in gain would explain the observed damping of the collective wave at lower electric fields and higher temperatures.

As pointed out in Sec. III, the major contribution to mixing collisions probably comes from successive absorption and emission processes. This combined rate is limited by the slower of these two processes. When the electron drift velocity is exactly v_s the rates of the two processes are expected to be equal. For higher drift velocities (i.e., higher external fields) the absorption rate, and therefore the mixing rate decreases. Higher temperatures tend to enlarge the spread in the electron distribution and would bring the absorption and emission rates closer together. Lower temperatures, on the other hand, tend to make the electron distribution more compact, which would decrease the absorption rate relative to the emission rate. The rate of mixing collisions then has opposite temperature and voltage dependence than the net gain. This loss of coherence, or second viscosity damping, would account for the damping of the collective wave at high electric fields and low temperatures. The increase in velocity with onset of loss of coherence or second viscosity is also observed in second-sound propagation.[11] It comes about because for few mixing collisions a large part of the collective wave energy travels straight through as unhindered phonons. With no normal collisions the energy transfer velocity would be v_s.

Another cause for decrease in the observability of the collective waves at low temperature is the decrease in the energy and momentum content of the effective band. Figure 5 shows the total crystal momentum transferred to the entire band for various temperatures. It can be seen that the total magnitude of the entire collective wave decreases with temperature.

We expect the mixing collision rate to be comparable to the rate of Eq. (5.1). Therefore, the highest collective frequencies possible at room temperature would be roughly 10^8 cps. For frequencies of 10^6 cps and lower, the collective wavelength would be as large or larger than the size of readily available crystals. The experimentally observed wave was observed at a frequency of 10^7 cps.

ACKNOWLEDGMENTS

The author wishes to thank Drs. H. Kroger, R. Damon, H. Carleton, R. Newman, P. Auer, and Professor H. Ehrenreich for many enlightening discussions. The author is indebted to Dr. H. Kroger and Mrs. G. Scott for critical reading and help in the preparation of this manuscript. The author also wishes to thank the referee for calling attention to Ref. 7 and for a number of helpful comments.

[12] A. R. Hutson, Phys. Rev. Letters 4, 505 (1960).
[13] T. O. Woodruff and H. Ehrenreich, Phys. Rev. 123, 1553 (1961).
[14] H. Kroger (private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Optical Harmonic Generation in Single Crystal $BaTiO_3$

ROBERT C. MILLER
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 21 January 1964)

Second-harmonic generation (SHG) of the Nd-doped $CaWO_4$ optical maser beam in ferroelectric $BaTiO_3$ has been investigated. Each of the three nonlinear coefficients which determine the magnitude of the SHG have been measured from room temperature to near the Curie temperature. It is found that all three coefficients have the same temperature dependence as the spontaneous ferroelectric polarization. The symmetry of the nonlinear coefficients required for a dispersionless, lossless SHG mechanism, namely, $d_{15}=d_{31}$ for $BaTiO_3$, applies over the entire measured temperature range. The effects on SHG due to antiparallel ferroelectric domains are described. The circumstances under which random as well as special antiparallel domain arrays can produce either an enhancement or a degradation of the second-harmonic intensity are discussed. Because of the dependence of SHG on the domain structure, the determination of the nonlinear coefficients requires single-domain crystals. A method for preparing suitable $BaTiO_3$ single-domain crystals is described. These crystals are also useful for investigations of the characteristics of ferroelectric domain growth.

INTRODUCTION

QUANTITATIVE studies of the second-harmonic generation[1] (SHG) of optical maser beams in ferroelectric crystals are interesting for several reasons. All ferroelectric crystals[2] have optical, electrical, and other properties that undergo large changes with temperature so that one might also expect the efficiency of SHG in these crystals to vary with temperature in some pronounced manner. In addition, ferroelectric crystals have a domain structure that can be influenced with an external electric field; and, as will be discussed, the ferroelectric domain structure can have a maked effect on the production of the second harmonic.[3–7] Neither the domain effects nor the marked temperature dependence of SHG will occur with the usual nonferroelectric piezoelectric crystal. Ferroelectric barium titanate is, in many respects, a good choice as a material for an investigation of these nonlinear effects. For example, suitable crystals are readily available, the point group (C_{4v}) is simple, it is among the most efficient nonlinear materials known,[6] the Curie temperature (120°C) is easily attained, its optical properties are well suited to SHG with at least one of the commonly used optical masers, and its domain dynamics are quite well understood.[2]

From the point group of $BaTiO_3$, it can be shown that the second-order polarization $\bar{P}_{2\omega}$ has the form,[1,6]

$$\bar{P}_{2\omega}=\begin{bmatrix}0&0&0&0&d_{15}&0\\0&0&0&d_{15}&0&0\\d_{31}&d_{31}&d_{33}&0&0&0\end{bmatrix}\begin{bmatrix}E_x^2\\E_y^2\\E_z^2\\2E_yE_z\\2E_xE_z\\2E_xE_y\end{bmatrix}. \quad (1)$$

In this equation, $\bar{P}_{2\omega}$ is the dielectric polarization at twice the maser frequency ω, d_{ij} are the nonlinear coefficients, and E_i are the optical electric fields in the medium. Inspection of Eq. (1) readily shows that c-domain ferroelectric plates, i.e., plates with the ferroelectric direction (z axis) normal to the major surfaces, are not very useful for a study of SHG. One would like to use a-domain plates, i.e., samples with the ferroelectric direction in the plane of the major surfaces of the plate, so that with appropriate use polarizers and analyzers, one can determine each of the three nonlinear coefficients. There is some indication in the literature that a-domain samples are difficult to prepare; and, furthermore, that they may be unstable.[8] In addition, since SHG is influenced by antiparallel ferroelectric domains,[6] one must prepare a-domain samples which are free of antiparallel domains, i.e., they must be single domain crystals. A method for preparing samples suitable for these SHG studies will be described.

The a-domain samples prepared for the SHG studies are also useful for another type of investigation. With electrodes along two opposite edges so that electric fields can be applied along the crystal z axis, one can study the domain dynamics of polarization reversal, i.e., the nucleation and growth of the antiparallel domains. In recent years, studies[2] of domain dynamics through the direct observation of the domains have been largely restricted to investigations of the sidewise motion of the 180° domain walls—the boundaries between antiparallel domains. With the a-domain sam-

[1] For a review on the subject of SHG, see P. A. Franken and J. F. Ward, Rev. Mod. Phys. 35, 23 (1963).
[2] For a general review of the properties of $BaTiO_3$ and other ferroelectric crystals, the reader is referred to F. Jona and G. Shirane, *Ferroelectric Crystals* (The Macmillan Company, New York, 1962).
[3] R. C. Miller, quoted in Ref. 1.
[4] F. Brown, Bull. Am. Phys. Soc. 8, 62 (1963).
[5] J. van der Ziel and N. Bloembergen, Bull. Am. Phys. Soc. 8, 380 (1963), and N. Bloembergen, Proceedings of the International School of Physics "Enrico Fermi," Varenna, Como, Italy, 1963 (to be published).
[6] R. C. Miller, D. A. Kleinman, and A. Savage, Phys. Rev. Letters 11, 146 (1963).
[7] R. L. Himbarger and J. L. Bjorkstam, Appl. Phys. Letters 3, 109 (1963).
[8] P. H. Fang and W. S. Brower, J. Appl. Phys. 34, 1516 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ples, one can for the first time observe in detail directly with a microscope the forward growth (growth in the direction of the spontaneous polarization) of the antiparallel domains under nearly ideal conditions. These phenomena, however, will not be described in this paper.

This paper discusses an analysis of some of the effects of antiparallel ferroelectric domains on SHG, the preparation of single crystals of $BaTiO_3$ suitable for studies of SHG and the forward growth of antiparallel domains, the determination of all the nonlinear coefficients as a function of temperature, and a brief discussion of these and other related experimental results.

EFFECTS OF FERROELECTRIC DOMAINS ON SHG

The effects which will be considered are those due to antiparallel domains. Effects which arise with twinned crystals due to mixed a- and c-domain configurations will not be discussed. Barium titanate will be treated in some detail, however, the phenomena to be described are not specific to $BaTiO_3$—they will occur with all ferroelectric crystals.

Consider what happens when an optical maser beam is at normal incidence on a $BaTiO_3$ crystal platelet. The maser beam fundamental produces a second harmonic forced wave[9,10] in the crystal as indicated in Fig. 1. This wave, which will be expressed as

$$E_{2k_1} \propto l_c \exp i(2k_1x-2\omega t) \tag{2}$$

is tied to the fundamental maser light wave, $\exp i(k_1x-\omega t)$. In these expressions, ω is the frequency of the fundamental light wave, and k_1 the propagation constant in the medium. The quantity l_c is the coherence length,[1] which will be described in more detail in the next paragraph. The plane of polarization of the forced wave is determined by the nonlinear coefficient involved in the SHG process. For ease of presentation, the phenomenon to be described will be illustrated with the nonlinear coefficient d_{33} so that all polarizations, including the ferroelectric polarization, are in the same plane.

Boundary conditions imposed by Maxwell's equations, namely, the tangential components of E and H, must be continuous across the surface of the crystal, require that a free second-harmonic wave[9,10] also be present in the crystal. This wave, which is produced at the surface, is expressed by

$$E_{k_2} \propto -l_c \frac{n_1+1}{n_2+2} \exp i(k_2x-2\omega t), \tag{3}$$

where k_2 is the propagation constant in the medium appropriate to the second-harmonic frequency. The quantities n_1 and n_2 are the indices of refraction at the fundamental- and second-harmonic frequencies, respectively. A reflected second-harmonic wave of small amplitude is also generated but will not be discussed further.[9,10] Since dispersion between the fundamental- and second-harmonic frequencies is usually present, $2k_1$ will not in general be equal to k_2 so that the forced and free waves, which are initially out of phase by π at $x=0$, travel with different velocities. The second harmonic intensity I_2 in the crystal is the result of the interference between these two waves, and can be approximated by

$$I_2 \propto |E_{2k_1}+E_{k_2}|^2. \tag{4}$$

Since $(n_1+1)/(n_2+1)$ is nearly unity, I_2 at $x=0$ is essentially zero compared to the maximum value of I_2 which first occurs at x equal to one coherence length l_c where the two waves are in phase; and then provided no domain walls are present, at successive odd multiples of l_c. The minima occur at even multiples of l_c. This oscillating phenomenon is illustrated near the bottom of Fig. 1. The coherence length l_c is given by

$$l_c=\pi/(k_2-2k_1)=\lambda/4(n_2-n_1), \tag{5}$$

where λ is the free-space wavelength of the fundamental.

Now, consider what happens to these two second-harmonic waves when they traverse a 180° domain wall

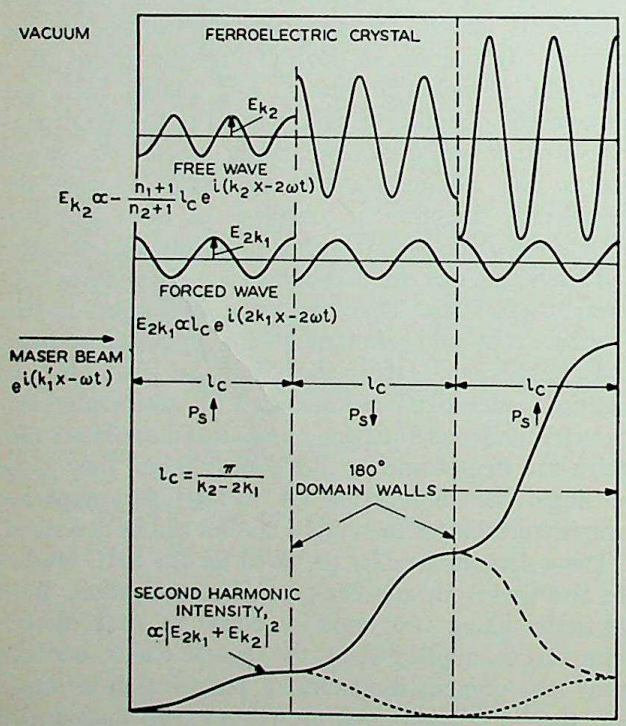


FIG. 1. Illustration of the enhancement of SHG due to antiparallel ferroelectric domains, one coherence length, l_c, thick. The forced and free second-harmonic waves are shown at the top of the figure and the resultant second-harmonic intensity at the bottom of the figure. At each domain wall the forced wave changes phase by π, and the free wave increases in amplitude by an amount equal to twice that of the forced wave. The dotted, dashed, and solid curves at the bottom of the figure show the effect on the second-harmonic intensity of zero, one, and two 180° domain walls.

[9] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. 127, 1918 (1962).
[10] D. A. Kleinman, Phys. Rev. 128, 1761 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where the normal to the wall is in the beam propagation direction. Since the z axes in two antiparallel domains are in opposite directions, Eq. (1) says that the phase of the second-order polarization wave $\bar{P}_{2\omega}$, and hence that of E_{2k_1}, will change by π on traversing the wall. The effect of the wall on the free wave is determined by the requirement that the tangential components of the E and H fields at frequency 2ω be continuous across the domain wall. For a wall at $x=x_0$, the free-wave changes from the form given in Eq. (3) to

$$-l_c\left[\frac{(n_1+1)}{(n_2+1)}-\frac{(n_1+n_2)}{n_2}\exp i(2k_1-k_2)x_0\right]$$
$$\times\exp i(k_2x-2\omega t), \quad (6)$$

for $x>x_0$. Thus, in the case of the free wave, the domain wall in general introduces a change in amplitude as well as a change in phase. A reflected second-harmonic wave is also produced at the wall; however, since its amplitude is down from that of the transmitted second-harmonic wave by more than three orders of magnitude, the reflected wave will not be considered further. Assume as shown in Fig. 1, that a wall is at $x_0=l_c$, in which case the maximum increase in E_{k_2}, and hence I_2 is obtained. Setting $n_1\approx n_2$, i.e., small dispersion, one sees from Eq. (6) that the amplitude of the free wave increases by a factor of 3. Therefore, the maxima of the total second-harmonic intensity increase by a factor of 4 due to the first "properly" placed domain wall. In addition, the minima are no longer zero, but are equal to the maxima obtained in the absence of the 180° wall. Each properly placed 180° wall results in an increase of the amplitude of the free wave by $(n_1+n_2)/n_2\approx 2$ times the amplitude of the forced wave. With N properly placed walls, the part of I_2 which does not average to zero with x can be shown to increase by a factor equal $[1+(1+2N)^2]/2$. The largest enhancement one can obtain from this process occurs when the 180° walls are present at $x=l_c$, $2l_c$, $3l_c$, etc., up to the crystal thickness l. The effect is then to increase the coherence length from l_c to the crystal thickness l. With the Nd maser and $BaTiO_3$, $l_c=2.1\times10^{-4}$ cm for d_{33}[6] so that for a crystal of typical thickness, 2×10^{-2} cm, and with domain walls spaced l_c apart, one has $N_{\max}=10^2$ or a SHG enhancement of 2×10^4, a very large effect.

For an array of N 180° domain walls situated randomly in the crystal, the resultant second-harmonic electric field is given by

$$E\propto l_c[(-1)^N e^{i2k_1x}-e^{ik_2x}(1-2\exp(-i\pi x_1/l_c)$$
$$+2\exp(-i\pi x_2/l_c)-2\exp(-i\pi x_3/l_c)+\cdots$$
$$+2(-1)^N\exp(-i\pi x_N/l_c))], \quad (7)$$

where the walls are placed at $x=x_1$, $x=x_2$, etc., and it is assumed that $n_1=n_2$ except in l_c. Then, for $2Nl_c<l$, Eq. (7) gives a second-harmonic intensity enhancement proportional to $(1+2N)$, where again only the part of I_2 that does not average to zero with x is considered. The restriction that $2Nl_c<l$ is required so that in averaging the product of Eq. (7) and its complex conjugate one can permit each phase $\pi x_i/l_c$ to vary from 0 to 2π. Thus, even a random array of antiparallel domains can produce a substantial enhancement of the second harmonic in a ferroelectric crystal.

For $l_c\gg l$, i.e., when the free and forced waves propagate at very nearly the same velocity, Eq. (7) can be used to show that the intensity of the resultant SHG is proportional to the square of the difference between the total thickness of domains of one sign minus that for domains of the opposite sign. Since SHG for $l\gg l_c$ in the absence of domains[1] is proportional to l^2, antiparallel domains will have a serious degrading effect on SHG under "velocity matched" conditions.

In the case just treated, the crystal was assumed to consist of slabs of antiparallel domains with the normals to the walls parallel to the beam direction. If the wall normals are perpendicular to the beam direction the analysis given above does not apply. In the latter situation, one has second-harmonic coherent light beams emerging from the crystal with the phase of all the beams coming from domains polarized in one direction out of phase by π with those beams coming from domains polarized in the opposite direction. In an unfocused beam experiment with a few antiparallel domains present so that the cross sections of the radiating volumes are much larger than a wavelength, the emerging beams will be nearly parallel and will not interfere with each other and reduce the second-harmonic intensity below that which one would observe with a single-domain crystal. However, when diffraction effects become large, that is, when the cross sections of the radiating areas are comparable to a wavelength, the emergent beams are no longer nearly parallel to that they can interfere destructively with each other and reduce the second-harmonic output below that which one would observe with one or a few domains.

The optimum domain array for SHG is one where the crystal consists of sheets of antiparallel domains, each l_c thick, with the wall normals parallel to the beam direction. The question arises as to how one can produce such domain arrays. One possible method that would apply to crystals which undergo second or higher order ferroelectric phase transitions,[2] such as potassium dihydrogen phosphate (KDP), Rochelle salt, and triglycine sulfate (TGS), involves the formation of a domain pattern, as the crystal goes from the paraelectric to the ferroelectric phase, which is determined by the minimum free energy.[2] In the case of Rochelle salt, theory and experiment give a domain spacing obtained in this manner which is proportional to the square root of crystal thickness.[11] (This crystal thickness refers to the direction of the ferroelectric polariza-

[11] T. Mitsui and J. Furuichi, Phys. Rev. 90, 193 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tion.) Thus, one may be able to produce a suitable array using this technique. This method will not work for crystals such as $BaTiO_3$ which undergo first-order phase transitions since the domain pattern formed on cooling through the phase transition is in this case determined by independent nuclei and not by the minimum free energy.[2]

It may be possible to produce desired domain arrays with electric fields. There is some evidence that the number of domains per unit area is determined by the magnitude of the field producing them,[12] so that with a suitable choice of the magnitude and the duration of the field, one may be able to produce useful domain patterns. A combination of electrical and thermal treatment may also be found useful.

Order of magnitude enhancements of the second-harmonic have been observed with both $BaTiO_3$ and TGS when antiparallel domains were known to be present. In these experiments, the ferroelectric domains had probably either circular, or approximately square, cross sections in the $BaTiO_3$; and probably lenticular, or circular, cross sections in TGS.[2] Thus, in neither case did the antiparallel domain configuration approximate the ideal slab structure. However, in no case was a decrease of the second harmonic observed when antiparallel domains were known to be present which shows that the enhancement effects of the 180° walls are larger than the degrading effects such as those due to small radiating areas.

A somewhat similar arrangement for enhancing SHG in quartz where $l_c < l$ was described earlier[1,9] and consists of stacking crystal plates in such a way so as to produce the effects shown in Fig. 1 and described in the text.

PREPARATION

Since no method of preparing samples suitable for the present experiments has been given in the literature, the techniques employed to produce the crystals will be described in some detail. Clear, undoped crystal plates of the order of 2×10^{-2} cm thick (grown by the Remeika method[13]) with clean smooth surfaces were selected and then etched in concentrated H_3PO_4 at 155°C sufficiently long to reduce the sample thickness by at least 5×10^{-3} cm. The samples were then rinsed in water, alcohol, and dc poled in distilled water with platinum electrodes. After poling, the specimens were examined with a polarizing microscope and only those samples which had large, unstrained c-domain areas were selected. Of these crystals, only those which had areas a millimeter or so on a side that were sufficiently plane parallel to give a few interference fringes with visible monochromatic light were used. These selected crystals were then cleaved, or broken, with a razor blade such that with much patience and some luck one obtained a c-domain sample, rectangular in shape, several millimeters on a side, with the sides parallel to the a axes. Then air-drying silver-paste electrodes were painted on the more perfect two opposite edges and dc fields of a few kV/cm were applied to pole the sample a domain. The poling process was monitored with a polarizing microscope and the process usually hastened by heating the sample with a focused light beam to a temperature below but near the Curie temperature. When the poling process was complete, the light and field were removed and the resultant a-domain sample examined. At this point, the samples would in most cases be without antiparallel domains. If a few antiparallel domains did appear, they were usually readily eliminated with fields of a few hundred volts per centimeter. In the studies of SHG, a field of this magnitude was always kept on the sample while under investigation to insure that no antiparallel domains were present. However, even under these conditions, both twinning and antiparallel domains usually occur at temperatures above approximately 110°C. For this reason, coupled with the fact that going through the Curie temperature is frequently a destructive operation, almost all the studies were made at temperatures less than about 110°C. There was no indication that these a-domain samples were unstable, i.e., over periods of months, no changes in the domain structure were observed.

SHG STUDIES

Figure 2 shows a schematic drawing of the essential components of the experimental arrangement for studying SHG in $BaTiO_3$. For ease of analysis of the data, a careful alignment was made so that the polarization of the laser beam was in the xz plane of the crystal and at 45° to the x and z crystalline axes. Table I gives the nonlinear coefficients d_{ij} involved in SHG as determined from Eq. (1) for various orientations of the polarizer and analyzer with respect to the crystal z axis. As seen in the table, special orientations of the polarizing elements enable one to observe separately SHG from each of the three nonlinear coefficients. The polarizing elements in Fig. 2 are set to observe SHG due to d_{15}.

To make quantitative comparisons between nonlinear

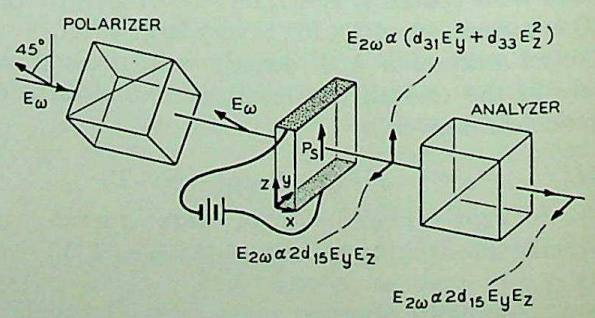


FIG. 2. Schematic drawing showing the orientation of the $BaTiO_3$ sample with respect to the maser beam. The polarizing elements are set to measure the second harmonic due to d_{15}. The electric field is applied to insure that the sample remains single domain.

[12] H. L. Stadler and P. J. Zachmanidis, J. Appl. Phys. 34, 3255 (1963).
[13] J. P. Remeika, J. Am. Chem. Soc. 76, 940 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

coefficients, it is necessary to consider the interference effects that occur between the second-harmonic forced wave produced by the maser beam, and the free second-harmonic light wave produced at the surface of the crystal. As mentioned in an earlier section, this interference effect, which was first demonstrated in quartz,[14] produces a second-harmonic wave whose amplitude varies in a periodic manner along the beam direction from essentially zero to some maximum value. For d_{33} in $BaTiO_3$ and a crystal of thickness l, the second-harmonic intensity, I_{33} generated by the fundamental intensity I_1 is given by

$$I_{33} \propto I_1^2 f(n) d_{33}^2 l_{33}^2 \sin^2(\pi l/2l_{33}), \tag{8}$$

where $f(n)$ is a function of the indices of refraction, and l_{33} is the coherence length. The function $f(n)$ is given by

$$f(n) = \frac{n_2}{(n_1+1)^3(n_2+1)^3(n_1+n_2)}. \tag{9}$$

In Eq. (8), the intensities are those measured outside of the crystal. Sufficient data on the indices of refraction of $BaTiO_3$ are not available to calculate the coherence length from Eq. (5) so that it must be determined experimentally from the data. The second-harmonic intensity is observed as the crystal is rotated about the z axis and then plotted as a function of angle. This rotation changes the path length of the beam in the crystal so that the periodic variation of the second harmonic is observed. These data on the angular dependence of I_{ij} are then used to determine l_{ij}, and are extrapolated to normal incidence to obtain relative values of d_{ij}. Room temperature determinations of d_{ij} and l_{ij} were given earlier,[6] but are also given here in Table II.

The temperature variation of the nonlinear coefficients is calculated from observations of the second harmonic intensity versus temperature. Figure 3 shows some of the data for the coefficient d_{31}. The oscillations present in the data occur due to changes in the magnitude of l_{31} through the temperature dependence of n_2-n_1.[6] Data on both the angular and temperature dependence of the second-harmonic intensity are required to determine directly the direction in which l_{ij} changes with temperature. For example, in the case of d_{15}, it is found that $dI_{15}/d\theta$ is negative when dI_{15}/dT is positive so that l_{15} must be increasing with temperature. Similar considerations show that l_{31} and l_{33} both decrease with temperature. These experimentally determined temperature dependences for l_{ij} are in the direction one would predict from the limited data on the optical properties of the negative uniaxial $BaTiO_3$ crystal. The published data[2] show that the index of refraction for the ordinary ray is temperature-independent while the index of refraction of the extraordinary ray increases with temperature and becomes equal to the ordinary index of refraction at the transition temperature.

Since $f(n)$ in Eq. (9) is very nearly temperature-independent, the temperature variation of the amplitude of the maxima shown in Fig. 3 arises from changes in l_{31} and d_{31}. Since l_{31} is known from the room temperature data on the angular dependence of the second harmonic intensity, and each oscillation of the second harmonic shown in Fig. 3 represents the introduction of two additional coherence lengths, l_{31} can be determined as a function of temperature. The temperature variation of d_{31} can then be calculated from Eq. (8). Figure 4 shows data obtained in the manner just described on the temperature variation of each of the nonlinear coefficients for $BaTiO_3$. The relative values of d_{ij} were computed at the intensity maxima and are normalized so

TABLE I. Orientations of polarizer and analyzer to determine nonlinear coefficients in $BaTiO_3$.[a]

Polarizer	Analyzer	Nonlinear coefficient
0°	0°	d_{33}
90°	0°	d_{31}
90°	90°	none
45°	0°	d_{33} and d_{31}
45°	90°	d_{15}

[a] The angles are between the transmission direction of the electric field in the polarizing element and the z axis in the crystal.

TABLE II. Room temperature $BaTiO_3$ nonlinear coefficients and coherence lengths.

d_{ij}[a]	l_{ij}
$d_{15}=35\pm3$	$1.57\pm0.12\,\mu$
$d_{31}=37\pm3$	$2.90\pm0.14\,\mu$
$d_{33}=14\pm1$	$2.07\pm0.04\,\mu$

[a] The d_{ij} are relative to d_{15} for KDP where $d_{15}=1.00$.

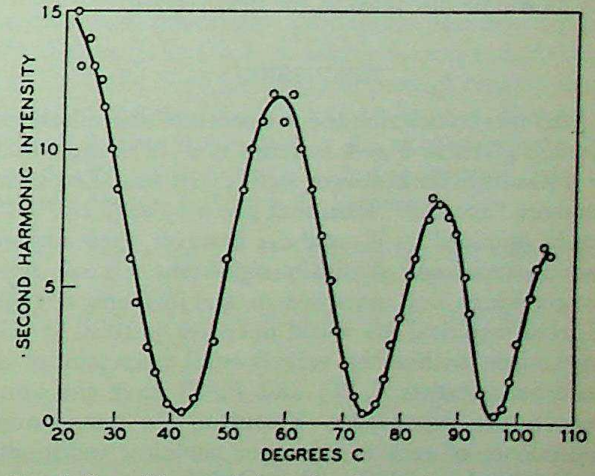


FIG. 3. Second harmonic intensity due to d_{31} as a function of temperature. The oscillations are due to changes in the indices of refraction with temperature. These data are used to calculate the temperature dependence of the nonlinear coefficient.

[14] P. D. Maker, R. W. Terhune, M. Nisenoff, and C. M. Savage, Phys. Rev. Letters 8, 21 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

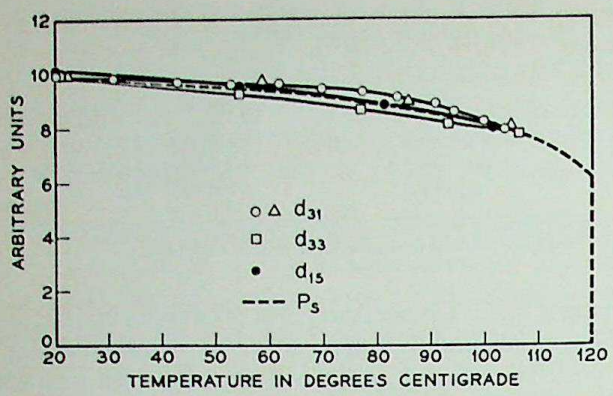


FIG. 4. The temperature dependence of the three nonlinear coefficients in $BaTiO_3$. All three have been normalized to 10 at 20°C. Within the experimental uncertainties of the measurements, all three coefficients vary with temperature as the spontaneous polarization which is also plotted in the figure and normalized to 10 at 20°C. The experimental points were calculated from the maxima of the second-harmonic intensity versus temperature data similar to that shown in Fig. 3.

that they all equal 10 at room temperature. All three coefficients were measured as a function of temperature with a crystal 2.48×10^{-2} cm thick, and in addition d_{31} was also measured with a second crystal 6.25×10^{-3} cm thick. In the case of the coefficient d_{15}, the curve shown in Fig. 4 is an average of six temperature runs. Although not shown in Fig. 4, it has been found that the d_{ij} decrease to zero discontinuously at the temperature where the crystal goes through a first-order phase transition from a ferroelectric phase into a nonpiezoelectric phase.[2] The discontinuous decrease of the second-harmonic intensity at the phase transition is consistent with the first-order phase transition. The temperature dependence of the spontaneous polarization of $BaTiO_3$ (26 $\mu C/cm^2$ at 300°K) normalized to 10 at room temperature, is also shown in Fig. 4. Note that all d_{ij} vary with temperature within a few percent of each other and P_s. The temperature dependence of the d's will be discussed in the next section.

DISCUSSION

Attempts to describe the temperature dependence of the d_{ij}'s given in Fig. 4 in terms of $P_s(T)n$ show that $n=1$ results in the best over-all fit to the data. Empirical fits more "accurate" than that given by $d_{ij}(T)\propto P_s(T)$ can be obtained for d_{31} and d_{33}, however, these fits are in some cases not physically significant, e.g., d_{ij} does not go to zero as P_s goes to zero, and furthermore complicated empirical fits would not seem justified at this stage since within the experimental uncertainties of the measurements $d_{ij}(T)$ and $P_s(T)$ have the same temperature dependence. Therefore, the temperature dependence of each of the three nonlinear coefficients for $BaTiO_3$ from 20°C to about 105°C can be described by

$$d_{ij}(T)=\alpha_{ij}P_s(T) \qquad (10)$$

in which $P_s(T)$ is the temperature-dependent spontaneous polarization, and α_{ij} is independent of temperature. The question arises as to the significance of Eq. (10). At first glance there may be some objection to trying to describe a purely optical property such as d_{ij} in terms of P_s which is a dc or at most a low-frequency characteristic. However, there is at least one purely optical quantity whose temperature dependence is described by some power of P_s. Namely, in $BaTiO_3$ the birefringence is proportional to P_s^2.[15] Also in KDP, the temperature dependence of the change in birefringence consequent on passing through the ferroelectric phase transition is proportional to P_s^2.[16] Therefore, at this point one can say that it may not be simply fortuitous that Eq. (10) describes the temperature dependence of the d_{ij}.

There are few other data with which one can compare the present results or test the general validity of Eq. (10). SHG studies of KDP both above and below the Curie temperature, 123°K, have been described in two recent publications.[5,7] Since SHG occurs in KDP both above and below the Curie temperature (it is piezoelectric in both phases) the temperature dependence of SHG in KDP will be different from that of $BaTiO_3$. The onset of ferroelectricity in KDP involves a phase change from the point symmetry D_{2d} for the paraelectric form, to the point symmetry C_{2v} for the ferroelectric phase. This lowering of the crystal symmetry results in the introduction of one nonlinear coefficient, d_{33}, not present in the paraelectric phase. Therefore, the temperature dependence of d_{33} will be different from that of the coefficients which are also present in the high-temperature phase. If it turns out that Eq. (10) does indeed have some general validity, it would be reasonable to propose that the temperature dependence of the nonlinear coefficients for a ferroelectric crystal which is piezoelectric in the paraelectric phase, such as KDP, is given by

$$d_{ij}(T)=d_{ij}^0+\alpha_{ij}P_s(T), \qquad (11)$$

where d_{ij}^0 is a temperature-independent nonlinear coefficient characteristic of the paraelectric phase. The change in the nonlinear coefficient in going through the phase change is described by $\alpha_{ij}P_s(T)$. For the new coefficient, d_{33}, which appears when the crystals becomes ferroelectric, $d_{33}^0=0$.

Himbarger and Bjorkstam[7] observe no significant change in the second-harmonic intensity from KDP in going from room temperature, through the phase transition, to 80°K. The data, which were obtained with a 0° ruby rod and therefore not a highly polarized maser beam, indicate that any discontinuity in the coefficients at the transition temperature is small. In terms of Eq. (11), these data would imply that $d_{ij}^0\gg\alpha_{ij}P_s$ except for d_{33} and also that $d_{33}\ll$ all other d_{ij}. The coefficient d_{33}

[15] W. J. Merz, Phys. Rev. 76, 1221 (1949). D. Meyerhofer, Phys. Rev. 112, 413 (1963).
[16] B. Zwicker and P. Scherrer, Helv. Phys. Acta 17, 346 (1944).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is reported by van der Ziel and Bloembergen[5] to be at least an order of magnitude smaller than the coefficients in the paraelectric phase. However, in disagreement with the data of Himbarger and Bjorkstam, van der Ziel and Bloembergen find that the nonlinear coefficients which are present in both phases change by factors ranging from about 3 to 0.2 as the crystal symmetry is lowered. Furthermore, except for the change in these coefficients at the phase transition, they are temperature-independent. The temperature dependence of d_{33} has not been determined. The discrepancy between the results of these two investigations on KDP could be due to effects arising from antiparallel ferroelectric domains; however, both groups were clearly aware of this complication. Discontinuities in the coefficients at the Curie temperature coupled with temperature-independent coefficients in the ferroelectric phase as reported by van der Ziel and Bloemberger are not consistent with Eq. (11).

Kleinman[17] has used a thermodynamic approach to show that if SHG arises from a nondispersive, nonlossy process, e.g., a high-frequency electronic mechanism, symmetry in addition to that required by the point group of the crystal will appear in the tensor which describes the second-order polarization coefficients. Data[6,18,19] on the Kleinman symmetry condition in a wide variety of crystals show that the additional symmetry is indeed present in the second-order polarization tensors. This symmetry condition requires for $BaTiO_3$ that $d_{15}=d_{31}$. As is evident in Table I, this condition is satisfied at room temperature within the approximately $\pm 10\%$ experimental uncertainties. The data shown in Fig. 4 demonstrate that the Kleinman symmetry condition is also satisfied over the entire measured temperature range.

A phenomenological theory[2] of ferroelectricity in which the Helmholtz free energy is expressed as a power series in the polarization has been quite successful, especially in the case of $BaTiO_3$ (Devonshire's theory), in explaining much experimental ferroelectric data. The validity of the Kleinman symmetry condition mentioned above, and the more general symmetry relations for nonlinear optical effects in lossless media derived using energy considerations by Pershan[20] both suggest that it may prove fruitful to extend the ferroelectric thermodynamic approach into the optical region. This approach would give $d_{ij}(T)$ expressed as a power series in P_s with the coefficients of the various terms determined from the free-energy expression for the paraelectric phase. For $BaTiO_3$ only odd powers of P_s will occur in d_{ij}, while d_{ij} for KDP will involve all powers of P_s. Therefore Eq. (10) includes only the first term for $BaTiO_3$, and Eq. (11) the first two terms for KDP. However, dispersion does occur between the present optical frequencies and the frequencies at which the various coefficients in the free-energy expression have been determined. Therefore, coefficients for the free-energy expression at optical frequencies as well as clarification of the KDP second harmonic data and/or additional SHG measurements with other ferroelectric crystals are required before the full significance of Eqs. (10) and (11) can be ascertained.

Another phenomenon which will be discussed briefly is the possible effect on the SHG due to a temperature dependence of the absorption edge in the $BaTiO_3$ crystals. As the crystal is warmed, the absorption edge, about 4000 Å at room temperature, shifts to slightly longer wavelengths.[21] It has been suggested[22,23] that the efficiency of SHG should increase as the second-harmonic wavelength approaches that of the absorption edge; however, to date, no data supporting this suggestion have been presented. In fact, data[6] on SHG in CdS show that d does not change significantly when the absorption edge is moved thermally through the frequency of the second harmonic. In any event, this edge effect would if anything give an increasing d_{ij} with temperature, which is opposite to what is observed. Therefore, if the proximity of the edge to the second-harmonic frequency is affecting the temperature dependence of the present data, it is being dominated by a still larger effect which results in d_{ij} which decrease with increasing temperature.

ACKNOWLEDGMENTS

The author would like to thank Dr. D. A. Kleinman for numerous informative discussions on many aspects of this paper; and Dr. J. A. Giordmaine, Dr. A. Ashkin, and Dr. G. D. Boyd for suggestions which resulted in an improved manuscript. The assistance of A. Savage in obtaining some of the data is also acknowledged. The fine $BaTiO_3$ crystals which made this research possible were kindly supplied by J. P. Remeika.

[17] D. A. Kleinman, Phys. Rev. **126**, 1977 (1962).
[18] A. Savage and R. C. Miller, Appl. Opt. **1**, 661 (1962).
[19] R. C. Miller, Phys. Rev. **131**, 95 (1963).
[20] P. S. Pershan, Phys. Rev. **130**, 919 (1963).
[21] R. C. Casella and S. P. Keller, Phys. Rev. **116**, 1469 (1959).
[22] B. Lax, J. G. Mavroides, and D. F. Edwards, Phys. Rev. Letters **8**, 166 (1962).
[23] R. Loudon, Proc. Phys. Soc. (London) **80**, 952 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Specific Heat of Niobium between 0.4 and 4.2°K*

B. J. C. VAN DER HOEVEN, JR., AND P. H. KEESOM
Purdue University, Lafayette, Indiana

(Received 26 December 1963)

Below 3.3°K, the normal-state specific heat of niobium, measured in a magnetic field of 17 kG, can be represented by $C_n = 7.79T + 0.094T^3$ mJ/mole deg. This corresponds to a value of the Debye parameter at 0°K, $\Theta(0)$, of 275°K. The specific heat in the superconducting state does not appear to be anomalous, as has been reported.

INTRODUCTION

THEORY predicts that the lattice specific heat of a superconductor should be the same in both the normal and superconducting states. The normal-state specific heat is a sum of an electronic term γT and the lattice term C_{ln}. Since for sufficiently low temperatures the superconducting electronic contribution C_{es} becomes negligible, the total superconducting-state specific heat C_s is equal to the lattice contribution C_{ls}. So, for reduced temperatures above $T_c/T = 5$ one expects $C_s = C_{ln}$.

Two independent measurements of indium have provided an exception to this rule, however.[1,2] The superconducting specific heat C_s was found to be 10–20% less than the normal lattice contribution C_{ln}. A similar anomaly was reported by Hirshfeld *et al.* for niobium.[3] Recent conversations with Dr. Boorse have encouraged independent measurements and he was kind enough to offer one of his samples. The results are reported below.

EXPERIMENT

The apparatus was the He^3 cryostat described previously.[4] A germanium thermometer, similar in composition to the one described in an earlier publication[5] was used. Resistance range of the thermometer was 40Ω at 4.2°K, 120Ω at 1.0°K, and 500Ω at 0.4°K. The thermometer was calibrated in separate runs at zero field and at 17 kG. Magnetoresistance of the thermometer, $\Delta\rho/\rho$, varied from minus 2% at 4.2°K to plus 8% at 0.4°K. The 1962-He^3 temperature scale of Sydoriak and Roberts[6] for the range 0.4 to 1.2°K was used in conjunction with a paramagnetic salt thermometer for values below 0.5°K for the zero-field measurements and the 1958-He^4 temperature scale[7] above 1.0°K. An expansion of $1/T$ in powers of $\ln R$ up to $(\ln R)^3$ was done on an IBM-7090 digital computer, using the method of least squares as described by Moody and Rhodes.[8]

Total correction for the addenda (thermometer, heater, and glyptal) was less than 10% of the total heat capacity of the sample in the superconducting state at the lowest temperature. As the heat capacity of the addenda is known within ±3%, this introduces an error of at most 0.3% in the specific heat. Use was made of the most recent data of Ho *et al.*[9] for the constantan heater wire.

Normal-state measurements were made with a new set of cans incorporating the use of a superconducting Nb_3Zr wire solenoid with an inner diameter of 2.5 cm. The working volume was 1.5 cm in diameter and 3.5 cm long. A field of 17 kG was applied. In the superconducting state, a Helmholtz coil arrangement was employed to cancel the earth's magnetic field. The cylindrical sample of 0.365 moles was the same one used by Hirshfeld *et al.* with a small piece cut from the bottom. Dimensions were 1.3 cm in diameter, and

TABLE I. Specific heat of niobium. (C is in mJ/mole deg.)

T (°K)	C_s	T (°K)	C_n
0.372	0.038	0.568	4.52
0.723	0.077	0.647	5.12
0.799	0.089	0.736	5.76
0.849	0.097	0.824	6.42
0.906	0.107	0.946	7.39
1.016	0.128	1.011	7.94
1.152	0.156	1.186	9.39
1.272	0.242	1.281	10.2
1.361	0.297	1.370	10.9
1.499	0.406	1.519	12.2
1.635	0.562	1.646	13.2
1.794	0.822	1.771	14.3
1.988	1.28	2.016	16.5
2.221	2.14	2.225	18.4
2.456	3.40	2.451	20.4
2.676	5.01	2.734	23.2
2.998	8.18	3.011	25.9
3.260	11.4	3.311	29.2
3.562	16.1	3.616	33.0
3.784	20.0	3.752	34.8
4.014	25.2	3.995	38.3
4.304	32.6	4.298	42.8

* Work supported by U. S. Army Research Office.

[1] C. A. Bryant and P. H. Keesom, Phys. Rev. 123, 491 (1961).
[2] H. R. O'Neal, Ph.D. thesis, University of California, UCRL-10426, 1963 (unpublished.)
[3] A. T. Hirshfeld, H. A. Leupold, and H. A. Boorse, Phys. Rev. 127, 1501 (1962).
[4] G. M. Seidel and P. H. Keesom, Rev. Sci. Instr. 29, 606 (1958).
[5] B. J. C. van der Hoeven, Jr. and P. H. Keesom, Phys. Rev. 130, 1318 (1963).
[6] S. G. Sydoriak, T. R. Roberts, and R. H. Sherman, in Eighth International Congress on Low Temperature Physics, London, 1962, p. 297 (unpublished).
[7] F. G. Brickwedde, H. van Dijk, M. Durieux, J. R. Clement, and J. K. Logan, J. Res. Natl. Bur. Std. 64A, 1 (1960).
[8] D. E. Moody and P. Rhodes, Cryogenics 3, 77 (1963).
[9] J. C. Ho, H. R. O'Neal, and N. E. Phillips, Rev. Sci. Instr. 34, 782 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

length 3.5 cm. The estimated experimental accuracy is ±2%.

RESULTS

The results for the superconducting and normal states are listed in Table I. They are also plotted in the form C/T versus T^2 in Fig. 1, together with a few points from the smoothed curve of the data of Hirshfeld *et al.* The two measurements are in very good agreement within combined experimental error. Up to 3.3°K the normal state data may be represented by

$$C_n = 7.79T + 0.094T^3 \text{ mJ/mole deg.}$$

The coefficient of the T^3 term α corresponds to a Debye parameter at 0°K, $\Theta(0)$, equal to 275°K.

In Fig. 2 is plotted an expanded portion of C_s/T versus T^2 up to $T^2=4$. If one plots here the values of

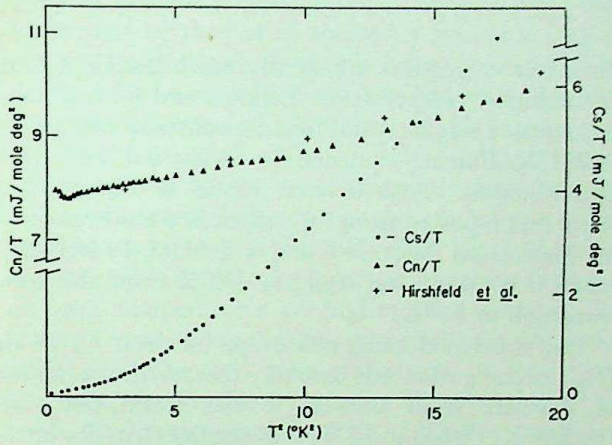


FIG. 1. C_s/T and C_n/T versus T^2.

C_s/T found by Hirshfeld *et al.*, no difference is observed. Two slopes are drawn representing the $\Theta(0)$ value of these measurements and the same quantity given by Hirshfeld *et al.* of 238°K. The small positive intercept of C_s/T (0.022 mJ/mole deg²) will be discussed later. It is clear from this graph that the anomaly disappears using the new value of $\Theta(0)=275$°K.

The data for $C_{es}/\gamma T_c$ can be very well represented by $a\cdot\exp(-bT_c/T)$. Plotting in this manner yields values of $a=8.31$ and $b=1.53$, very close to those predicted by the BCS theory of $a=8.5$ and $b=1.44$.

DISCUSSION

The lattice specific heat of a metal may be expressed as

$$C_l = 1944(T/\Theta(T))^3 \text{ J/mole deg},$$

where the Debye parameter is a function of tempera-

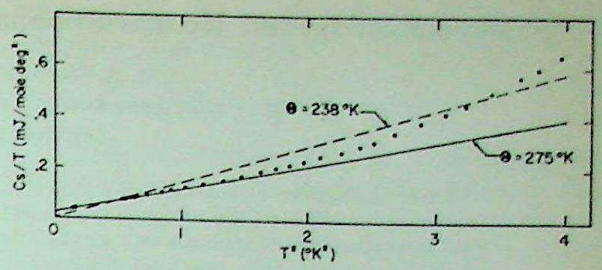


FIG. 2. C_s/T versus T^2.

ture. As mentioned before, the agreement between the values for the specific heat as measured by Hirshfeld *et al.* and these measurements are in excellent agreement within experimental error. However, the additional points at lower temperatures and with higher magnetic field modify the interpretation of the results. The initial value of Θ is now 275°K, in excellent agreement with the value derived from velocity of sound measurements by Alers and Waldorf of 277°K.[10] At 3.3°K, the value of the Debye parameter gradually decreases to a value of 238°K above 4.5°K, indicating the presence of a T^5 or higher term in the specific heat.

The increase of the normal-state specific heat at the lowest temperature may be a latent heat contribution due to insufficient quenching of the superconductivity. However, the magnitude of this deviation is within our experimental accuracy and could be due to inaccuracies in the thermometer calibration curve.

In Fig. 2 one observes that the specific heat points C_s/T gradually become tangent to the slope corresponding to $\Theta=275$°K at 1°K. The points do not fall below this line. Thus, the condition that $C_s=C_{ln}$ is met and no anomaly is observed.

The positive intercept of this curve with the C_s/T axis of 0.022 mJ/mole deg² indicates a small contribution of normal electrons. Comparing this to the value of γ in the normal state, one deduces that 0.3% of the niobium was normal. This effect may be attributed to strains or impurities. Trapping of the earth's magnetic field would make a contribution an order of magnitude too small. However, the lattice specific heat should not be influenced by this small contribution.

In conclusion it appears unnecessary to assume that the specific heat of superconducting niobium is anomalous.

ACKNOWLEDGMENT

The authors are indebted to Dr. Boorse who put his sample at their disposal.

[10] G. A. Alers and D. L. Waldorf, Phys. Rev. Letters 6, 677 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Superconducting and Normal Specific Heats of a Single Crystal of Niobium*

H. A. LEUPOLD AND H. A. BOORSE†
Pupin Physics Laboratories, Columbia University, New York, New York

(Received 24 January 1964)

The anomalous specific heat found in a polycrystalline sample of niobium and previously reported from this laboratory by Hirshfeld, Leupold, and Boorse (HLB) prompted complete remeasurements both in the normal and superconducting phases on a single crystal with lower tantalum content. The thermal constants of the single crystal show only small differences from the polycrystal in the same temperature range, the single-crystal values being $T_c=9.20$°K and $\Theta=241$°K between about 10° and 3°K. The normal-state single-crystal measurements agree with measurements on the HLB polycrystal made both in this laboratory and by van der Hoeven and Keesom at Purdue University. The latter measurements were carried to about 0.4°K in a field of 17 kG. As a result of their investigation, they report a Debye Θ of 275°K in the region below 3°K. With this value of Θ, the anomaly in the specific heat disappears. As both the single-crystal results and the van der Hoeven and Keesom values agree in the mutually measured ranges, Θ is taken to be 275°K. The corresponding value of γ is found to be 7.80 mJ/mole deg^2. The usual exponential behavior of the electronic specific heat $C_{es}/\gamma T_c=ae^{-bT_c/T}$ with $a=8.21$ and $b=1.52$, is observed over a restricted temperature range. Below $t=5$ the data exhibit larger values than predicted by the exponential in agreement with a number of other superconductors. The value of the energy gap at 0°K was found to be 3.69 kT_c. Overall comparisons are made with the BCS theory and a modification due to Swihart.

I. INTRODUCTION

EXTENSIVE measurements of the specific heat of niobium in the temperature interval from approximately 12 to 1.2°K have been reported previously from this laboratory.[1–3] The latest measurements made by Hirshfeld, Leupold, and Boorse[3] (HLB) on a polycrystalline specimen showed that the normal phase, obtained by the application of a magnetic field of approximately 10 000 G, had a specific heat which could be represented by the usual sum of an electronic and lattice term, viz., $C_n=\gamma T+\text{const}(T/\Theta)^3$ with $\gamma=7.53$ mJ/mole deg.[2] The data indicated that in the range 12 to about 3°K, the value of Θ was constant and equal to 238°K. At lower temperatures the specific-heat values were larger than those given by the C_n relation, but the presence of 0.2% tantalum in the sample and the very high fields[4] required to suppress completely superconductivity in Nb alloys prompted the explanation of the larger values as a consequence of the appearance of the mixed phase in the material. Accordingly, the value of Θ appropriate to the normal phase over the whole interval to absolute zero was taken as constant and equal to 238°K. Previous measurements on annealed Nb made by Chou, White, and Johnston[5] gave values varying from about 250°K near the transition temperature to approximately 320 near 2°K. The precision of the latter values was apparently much less than those at the higher temperatures. Leupold and Boorse[2] using an annealed single crystal found a constant value equal to 241°K. Further evidence for a constant value of Θ approximately equal to that found in the previous researches reported from this laboratory has been given by Morin and Maita.[6] Using a field of 18 kG, they found Θ constant and equal to 230°K from above the transition to 1.5°K.

The superconducting-phase specific heat C_s in the HLB investigation was carefully measured in a number of separate trials because it was found that with $\Theta=238$°K, C_s below 1.7°K was smaller than the lattice specific heat in the normal phase. Although a similar anomaly had previously been found in indium by Bryant and Keesom,[7] the fact that Nb is a hard superconductor and likely to show significant variation in properties from sample to sample, prompted further measurements on a better specimen. Preliminary measurements[2] on a single-crystal rod, $\frac{3}{8}$ in. in diameter, obtained from Materials Research Corporation, Orangeburg, New York, did not show the anomaly. Accordingly, careful measurements were made again on both the polycrystal and on another piece of the same single-crystal rod mentioned above. The remeasurement of the polycrystal yielded identical results within the error of measurement with the originally published values.[3] In order to improve the superconducting properties of the second piece of the single-crystal rod, it was subjected to a zone-refining process. This operation was effected through the kindness of Dr. J. E. Kunzler and E. Buehler of the Bell Telephone Laboratories, Murray Hill, New Jersey, where the specimen was subjected to four passes in an electron-beam melting apparatus in high vacuum. The specimen, as subsequently analyzed, showed that as compared to the HLB polycrystalline

* This research was supported in part by the National Science Foundation, the U. S. Office of Naval Research, and the Linde Air Products Company.

† Permanent address: Barnard College, Columbia University, New York, New York.

[1] A. Brown, M. W. Zemansky, and H. A. Boorse, Phys. Rev. **92**, 52 (1953).

[2] H. A. Boorse and H. A. Leupold, 1961 (unpublished).

[3] A. T. Hirshfeld, H. A. Leupold, and H. A. Boorse, Phys. Rev. **127**, 1501 (1962).

[4] R. M. Bozorth, A. J. Williams, and D. D. Davis, Phys. Rev. Letters **5**, 148 (1960).

[5] C. Chou, D. White, and H. L. Johnston, Phys. Rev. **109**, 788 (1958).

[6] F. K. J. Morin and J. P. Maita, Phys. Rev. **129**, 1115 (1963).

[7] C. A. Bryant and P. H. Keesom, Phys. Rev. **123**, 491 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

material,[3] the tantalum content was lower, being 750 parts per million as compared with 2000; N_2, 54 ppm as compared with 248; O_2 at 34 ppm, Fe at 25 ppm and Si at 75 ppm were all slightly higher. Otherwise, the analysis was the same. A series of x-ray photographs taken at intervals along the length of the specimen showed identical crystal structure.

II. APPARATUS AND PROCEDURE

The apparatus used to make the measurements on this single crystal was essentially the same as that used in the HLB investigation.[3] The principal change was the replacement of the inner He-4 bath in the cryostat with an He-3 reservoir in order to reach lower temperatures. In carrying out the investigation, additional precautions were taken to insure the reliability of the temperature measurements.

As in previous investigations, temperatures were determined by the use of resistance thermometers. To measure temperatures in the He-3 range, Allen-Bradley 0.1-W carbon composition resistors having a room temperature resistance of approximately 10 Ω were used. For the He-4 range and up to 20°K, Allen-Bradley 0.1-W resistors having an approximate room temperature resistance of 33 Ω were employed. After each run, a resistance-vapor pressure calibration of the thermometer was made so that any errors arising from cycling to nitrogen temperatures or higher were avoided. The vapor-pressure temperature conversions for He-3 were made using the 1962 scale of Roberts, Sydoriak, and Sherman,[8] for the He-4 region of the 1960 scale,[9] and for the hydrogen region the scale of Wooley, Scott, and Brickwedde.[10]

The resistance-temperature curve for the He-3 region joined smoothly to the curve for the He-4 region and in the range of overlapping temperatures gave indistinguishable values. A least-squares fit of the helium calibration date was obtained for each run with the aid of an IBM 1620 computer using the relation

$$T=\frac{B}{\ln R+K/\ln R-A}.$$

The temperatures calculated by means of this formula agreed with those determined experimentally from vapor pressure measurements to within 2 or 3 millidegrees.

For each run in which heat capacity data were taken above the helium temperature region, the thermometer was also calibrated against the vapor pressure of liquid hydrogen in the range between its normal boiling point (20.4°K) and triple point (13.957°K). The H_2 and He^4 points together were then fitted with the aid of a 1620 computer to the expression

$$\ln R=A+BT^{-1/2}+CT^{-1}.$$

The agreement between values of T calculated from this formula and those obtained by observed vapor pressures was excellent, the root-mean-square error amounting to about 6 millidegrees over both the helium and hydrogen ranges. This formula was used only in the specific heat calculation at the temperatures in the gap between the He and H_2 regions where a 6-millidegree error is only of the order of 0.1% of the measured temperature.

In the uncalibrated region between liquid helium and liquid hydrogen temperatures, it was essential that some checks be provided in order that the accuracy of the specific-heat determinations as well as that of the transition temperature could be estimated. To make such checks, gas thermometer measurements of the transition temperature of a highly purified lead wire were used. The gas thermometer measurements were made at Rutgers University through the kindness of Professor B. Serin. Two determinations of the transition temperature of the lead wire were made, yielding values of 7.210±0.005° and 7.209±0.005°K. Throughout the heat-capacity measurements, the lead wire was thermally bonded to the niobium sample and readings of the resistance thermometer on the sample determined when the lead changed from the superconducting to the normal phase or vice versa. In three separate comparisons, the reading of the resistance thermometer when placed in the $\ln R$ versus T formula above yielded the following temperatures: 7.19, 7.21, and 7.21°K. It was therefore concluded that the formula based on calibrations in the He-4 and hydrogen temperature ranged was not in error at 7.21°K by more than a few tenths percent. The adoption of 7.21°K as the lead transition

TABLE I. Smoothed values of the molar specific heat of niobium.

T (°K)	T^2 (°K)^2	C_e/T (mJ/mole deg^2)	C_n/T (mJ/mole deg^2)	T_c/T
0.0	0.00	0.000	7.80	∞
0.5	0.25	0.023	7.82	18.39
1.0	1.00	0.101	7.89	9.19
1.5	2.25	0.266	8.00	6.13
2.0	4.00	0.647	8.17	4.60
2.5	6.25	1.450	8.38	3.68
3.0	9.00	2.683	8.64	3.06
3.5	12.25	4.374	9.09	2.63
4.0	16.00	6.294	9.61	2.30
4.5	20.25	8.289	10.20	2.04
5.0	25.00	10.42	10.87	1.84
5.5	30.25	12.78	11.60	1.67
6.0	36.00	15.30	12.40	1.53
6.5	42.25	17.93	13.27	1.41
7.0	49.00	20.66	14.22	1.31
7.5	56.25	23.54	15.23	1.23
8.0	64.00	26.48	16.31	1.15
8.5	72.25	29.49	17.46	1.08
9.0	81.00	32.53	18.68	1.02

[8] R. H. Sherman, S. G. Sydoriak, and T. R. Roberts, LAMS—2701, 1962 (unpublished).
[9] F. G. Brickwedde, H. van Dyke, M. Durieux, J. R. Clement, and J. K. Logan, J. Res. Natl. Bur. Std. (U. S.) 64A, 1 (1960).
[10] H. W. Wooley, R. B. Scott, and F. G. Brickwedde, J. Res. Natl. Bur. Std. (U. S.) 41, 379 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

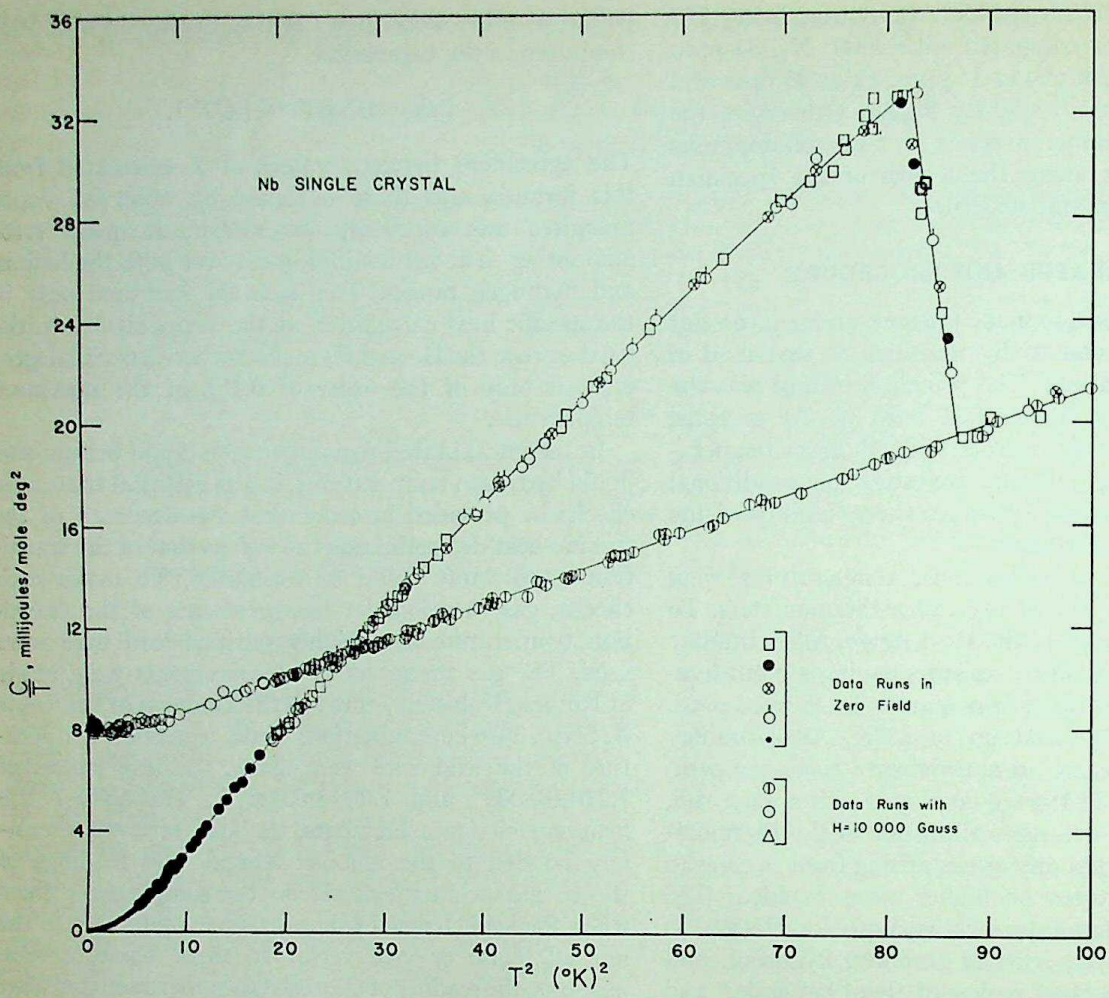


FIG. 1. C/T versus T^2 plot of the normal and superconducting specific heats of a single crystal of Nb. Full lines represent the mean values of the data.

temperature is 0.02°K higher than the latest determination[11] but more in line with the average of earlier measurements.[12]

III. RESULTS

The superconducting specific heat measured in zero field and the normal specific heat measured in a field of 10 000 G are shown in Fig. 1 and smoothed values for the data in the two phases are given in Table I. The normal data from the transition down to about 3°K can be calculated from the usual Debye-Sommerfeld relation with $\Theta = 241$°K and $\gamma = 7.38$ mJ/mole deg^2 obtained from the extrapolation of the $\Theta = 241$°K line to 0°K in substantial agreement with the values previously reported for the polycrystal. In the present data, however, starting at T^2 values between 10 and 9, the experimental points at decreasing temperatures rise gradually above the $\Theta = 241$°K normal line and fall on a new line whose Θ value was found to be 275°K with the corresponding value of γ equal to 7.80 mJ/mole deg^2. The uncertainty regarding the proper interpretation of this line suggested that the normal specific heat be determined in fields higher than the 10 000-G field available. Professor P. H. Keesom of Purdue University kindly agreed to measure the specific heat of the polycrystalline sample in a field of 17 kG in order to check the published data. The results of this investigation conducted by van der Hoeven and Keesom[13] (VK) and

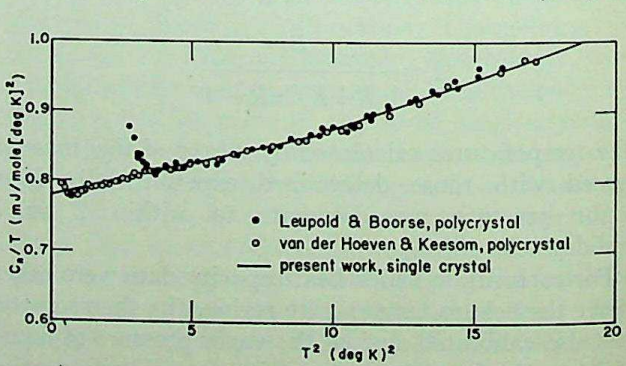


FIG. 2. C/T versus T^2 plot showing the experimental determinations of the normal specific heat below $T^2 = 20$. The data of van der Hoeven and Keesom is compared with polycrystal measurements in this laboratory, and the present data on a single crystal (full line) found in this investigation. The change in slope of the normal line just below $T^2 = 10$ should be noted.

[11] J. P. Franck and D. L. Martin, Can. J. Phys. 39, 1320 (1961).

[12] D. B. Cook, M. W. Zemansky, and H. A. Boorse, Phys. Rev. 78, 635 (1950).

[13] J. C. van der Hoeven and P. H. Keesom, Phys. Rev. (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

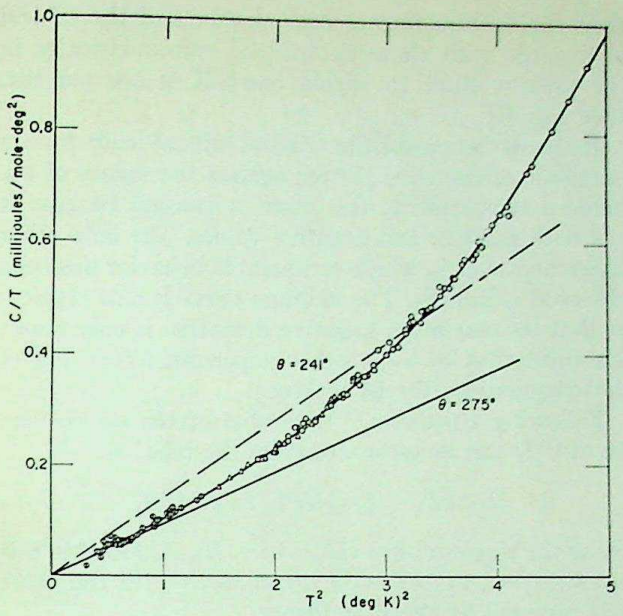


FIG. 3. C/T versus T^2 plot of the superconducting specific heat below $T^2=5$. The lines for Debye Θ values of 241° and 275° are shown for comparison.

published in this journal show clearly that the Θ = 275°K line is characteristic of the normal phase of niobium at the lowest temperatures and is not the result of the appearance of the superconducting phase.

In Fig. 2, the VK measurements are compared with those made in this laboratory on the same sample as well as with the smooth curve of the single-crystal measurements. All three sets of measurements agree within 1%. The measurements of VK show the same rise at 3° and follow the line Θ = 275 to about 0.7°K before undergoing an abrupt upward departure in the 17 000-G field. As can be seen from Fig. 2, the measurements made in this laboratory on the sample in a 10 000-G field rise sharply at about 1.8°K. This suggests that the gradual departure between $T^2=10$ and $T^2=9$ of the normal data from the line Θ = 241°K to the line Θ = 275°K reflects an inherent property of the crystal structure and not a consequence of an insufficient magnetic field as had been originally supposed. On this basis, the previously reported specific-heat anomaly disappears and the slope of the normal line corresponding to a Debye Θ of 275°K agrees with the limiting slope of the superconducting data close to the origin (see Fig. 3). It is interesting also that this value for Θ now agrees with the 277°K value found by Alers and Waldorf[14] from ultrasonic measurements at 4.2°K.

The transition temperature of the single-crystal sample was taken to be 9.19₅°K, the midpoint of the superconducting specific-heat discontinuity which extended from 9.07 to 9.32°K. A further confirmation for the view that the phonon spectrum of the metal changes in the neighborhood of 3°K is found in the experimentally determined entropy difference between the normal and superconducting phases. A plot of this difference versus temperature was found to vanish at 9.19₃°K. This value is in excellent agreement with the transition temperature taken as defined above, since the difference between 9.193° and 9.195° produces a disagreement in entropy between normal and superconducting phases of only 0.1%, less than the experimental error. If it is assumed that the change in Θ is an effect arising from imperfectly quenched superconductivity, the entropy difference vanishes at 9.13°K with the result that at 9.19₅°K the superconducting entropy exceeds the normal by about 3%. Thus an increase in Debye Θ seems to be the correct explanation for the change in slope of the normal line.

For a thermodynamically reversible superconductor, the critical field H_T at temperature T may be calculated from the relation

$$\int_T^{T_c}(S_n-S_s)dT=-(VH_T^2/8\pi). \qquad (1)$$

The entropy difference (S_n-S_s) for the single crystal was found by plotting a large scale graph of $(C_n/T-C_s/T)$ against the temperature and then measuring the appropriate areas with a planimeter. Using the values for the entropy difference thus obtained, Eq. (1) was integrated graphically and H_T could then be calculated after choosing $V=10.8$ cc/mole.[15] The resulting temperature dependence of H_T is shown on the graph in Fig. 4 and numerical values of H_T are given in Table II. The critical field H_0 at 0°K was found to

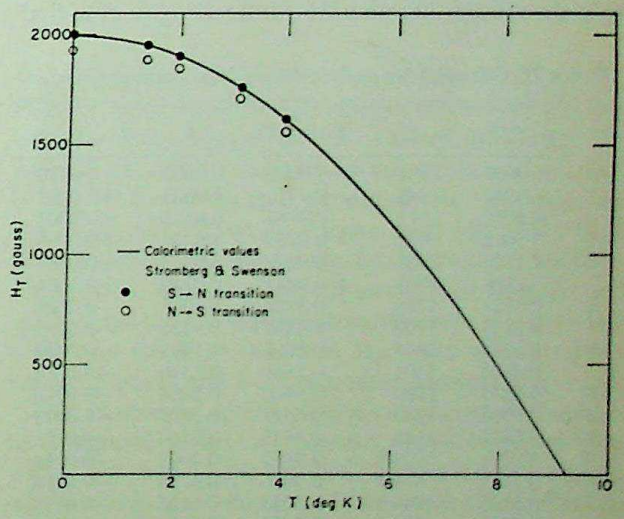


FIG. 4. Critical field versus temperature for the single-crystal specimen of Nb deduced from the specific-heat measurements, compared with critical-field measurements taken from the data of Stromberg and Swenson for transitions from normal to superconducting (open circles) and superconducting to normal (dots). The agreement of the latter data with the critical field deduced from these specific-heat measurements is obvious.

[14] G. A. Alers and D. L. Waldorf, Phys. Rev. Letters 6, 677 (1961).

[15] J. W. Edwards, R. Speiser, and H. L. Johnston, J. Appl. Phys. 22, 424 (1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

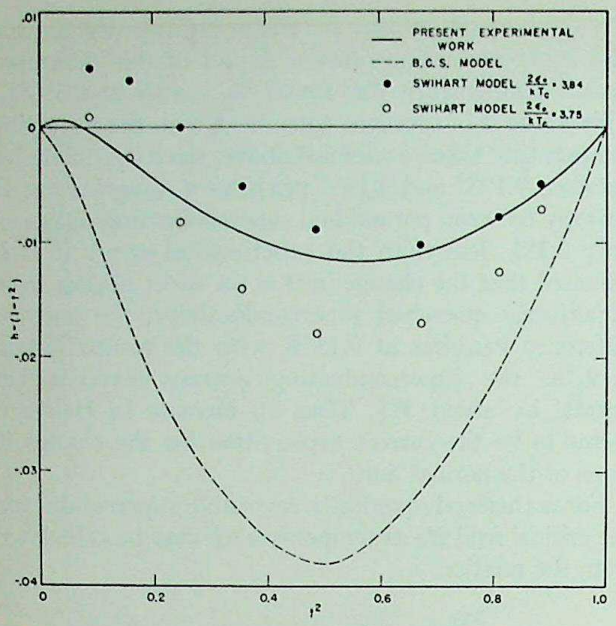


FIG. 5. Deviation of the critical field of the single-crystal Nb specimen determined from the specific-heat measurements from the parabolic dependence (full line). Dashed curve is the deviation according to the BCS theory. Open circles and dots are for comparison with Swihart's calculations.

be 1994 G. This is in excellent agreement with the field deduced from calorimetric measurements by McConville and Serin[16] who found the value 1990 G. The value deduced from direct magnetic measurements by Stromberg and Swenson[17] was 1960±40 G. The critical-field values of the latter investigation were obtained by taking the average of values measured when going from the superconducting to the normal phase and when going in the reverse direction. It is interesting to note that their superconducting to normal values of the critical field agree with those determined calorimetrically in the present work to within one-half of one percent. (See Fig. 4.)

TABLE II. Critical fields and deviation from parabolic law.

T	$t=T/T_c$	H_T	$h=H_T/H_0$	t^2	$h-(1-t^2)$
0.0	0.0000	1994	1.0000	0.0000	0.0000
0.5	0.0544	1988	0.9972	0.0030	+0.0002
1.0	0.1088	1971	0.9986	0.0118	+0.0004
1.5	0.1632	1942	0.9770	0.0266	+0.0006
2.0	0.2176	1901	0.9532	0.0473	+0.0005
2.5	0.2720	1847	0.9262	0.0740	+0.0001
3.0	0.3263	1780	0.8926	0.1065	−0.0009
3.5	0.3807	1700	0.8527	0.1450	−0.0023
4.0	0.4351	1608	0.8067	0.1893	−0.0040
4.5	0.4895	1501	0.7546	0.2396	−0.0058
5.0	0.5439	1389	0.6966	0.2958	−0.0076
5.5	0.5983	1262	0.6329	0.3579	−0.0092
6.0	0.6527	1124	0.5635	0.4260	−0.0105
6.5	0.7071	974.3	0.4887	0.4999	−0.0114
7.0	0.7615	814.8	0.4086	0.5798	−0.0116
7.5	0.8158	644.6	0.3233	0.6656	−0.0111
8.0	0.8702	465.1	0.2333	0.7573	−0.0095
8.5	0.9426	275.6	0.1382	0.8549	−0.0069
9.0	0.9790	78.19	0.0392	0.9585	−0.0023

If the deviations of the reduced critical fields from a parabolic relation are plotted against the square of the reduced temperature, the curve is unusual because it has both positive and negative values. The only other superconductor in which comparable behavior has been observed is indium. The niobium curve is also atypical in that its maximum negative deviation is only about one-third that of most other superconductors and of that required by the BCS theory.

Following Goodman,[18] the value of the energy gap $2\epsilon_0$ at 0°K can be estimated from the relation:

$$2\epsilon_0/kT_c=(\tfrac{4}{3})\pi(H_0^2V/8\pi\gamma T_c^2)^{1/2}.$$

Using the single-crystal values of γ, H_0, T_c found in this investigation, $2\epsilon_0/kT_c=3.69$, compared with the value 3.52 required by the BCS theory.

The value obtained from ultrasonic measurements by Dobbs and Perez[19] depends on the propagation direction in the crystal; being 3.77 for the [100] direction, 3.74 for the [111], and 3.65 for the [110]. Levy, Kagiwada, and Rudnick,[20] using transverse waves, found 3.7 and 3.5, depending on the method of analysis. The value found from the tunneling experiments of Giaever[21] was 3.6, from Sherril and Edwards[22] 3.59, and from Townsend and Sutton[23] 3.84±0.06. A value of 3.8±0.2 has been found by Mendelssohn[24] from thermal conductivity measurements. The present value lies in the range of these determinations.

The BCS theory predicts an electronic specific heat at the transition C_{es} equal to $2.43\gamma T_c$. The value found for $C_{es}/\gamma T_c$ in these measurements is 2.87.

If the electronic specific heat C_{es} is calculated in the usual way, viz:

$$C_{es}=C_s-C_n+\gamma T;$$

using $\gamma=7.80$ mJ/mole deg^2, the plot of $\ln C_{es}/\gamma T_c$ versus T_c/T gives a straight line for values $5.5>T_c/T>1.7$, indicative of the relation

$$C_{es}/\gamma T_c=ae^{-bT_c/T}.$$

This is shown in Fig. 6. The data obtained from this plot yields $a=8.21$ and $b=1.52$ in fair agreement with the predicted BCS values of 8.5 and 1.44. At values of

[16] T. McConville and B. Serin, International Conference on the Science of Superconductivity, Colgate University, August 1963 (unpublished).
[17] T. F. Stromberg and C. A. Swenson, Phys. Rev. Letters 9, 370 (1962).
[18] B. B. Goodman, Compt. Rend. 246, 3031 (1958).
[19] E. R. Dobbs and J. M. Perez, International Conference on the Science of Superconductivity, Colgate University, August 1963 (unpublished).
[20] Moises Levy, Reynold Kagiwada, and Isidore Rudnick, Phys. Rev. 132, 2039 (1963).
[21] I. Giaever, *Eighth International Conference on Low Temperature Physics* (Butterworths Scientific Publications Ltd., London, 1962).
[22] M. D. Sherril and H. H. Edwards, Phys. Rev. Letters 6, 460 (1961).
[23] P. Townsend and J. Sutton, Phys. Rev. 128, 591 (1962).
[24] K. Mendelssohn, IBM J. Res. Develop. 6, 27 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

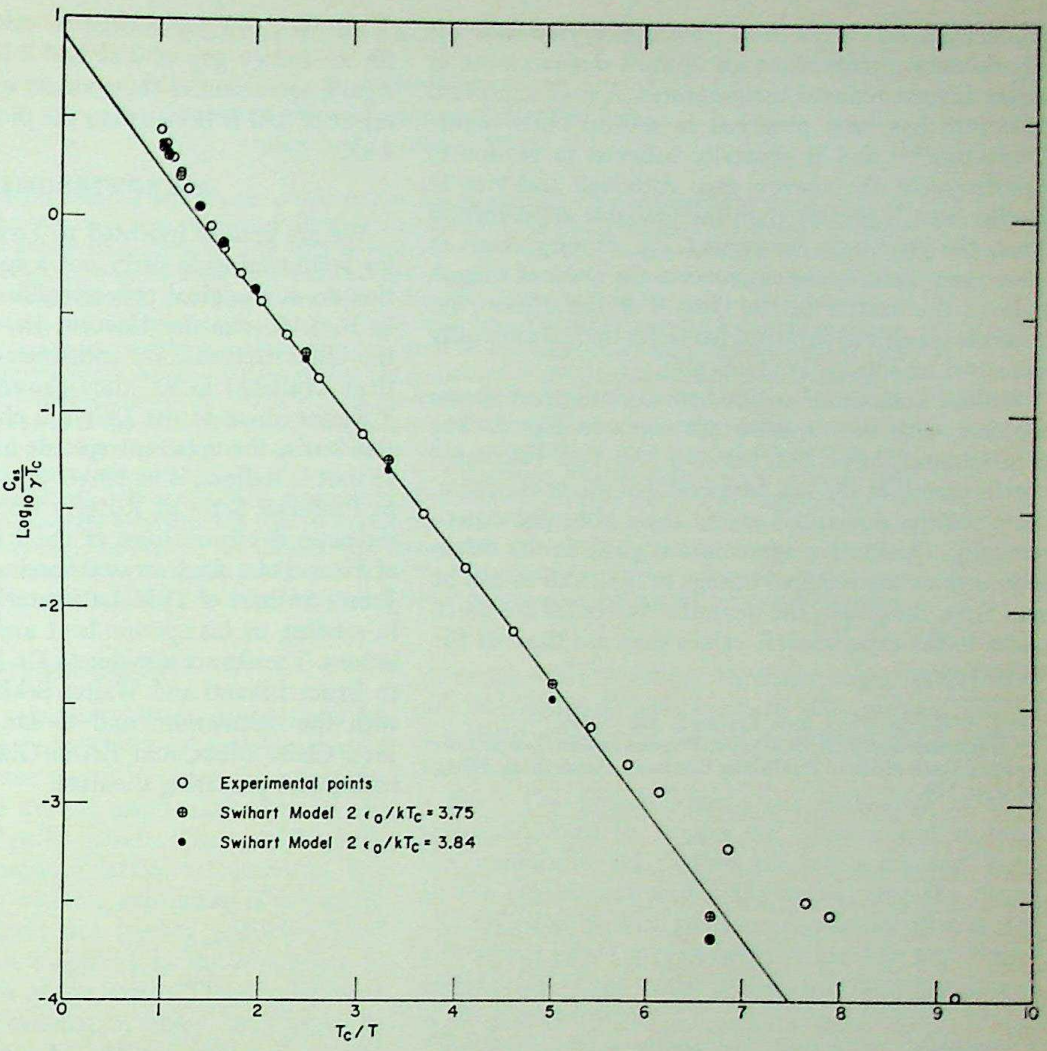


FIG. 6. The superconducting electronic specific heat C_{es} of a single-crystal specimen of Nb plotted on $\ln_{10}(C_{es}/\gamma T_c)$ versus T_c/T axes. Full line is the graph of $C_{es}/\gamma T_c = ae^{-bT_c/T}$, where $a=8.21$ and $b=1.52$. The dots and crossed circles show values computed from Swihart's model.

$T_c/T > 5.5$ the data departs from the exponential as will be discussed below.

Although the BCS theory gives fair agreement with the a and b values, the critical-field deviation curve, the value of the energy gap and the specific-heat jump at the transition all show significant disagreement. However, the BCS calculations were made under the assumption of weak coupling, i.e., that T_c/Θ is small, of the order of 0.003 or less, as for elements such as Zn, Al, and Ga. When T_c/Θ is an order of magnitude greater, however, as is the case for Nb (0.033 for $\Theta = 275°K$ and 0.039 for $\Theta = 241°K$) and In (0.039), the energy gap value increases to about 3.7 and the critical-field deviation has a minimum only about one-third to one-half as low. (See Fig. 5.) Moreover, the deviation actually is positive at the lowest temperatures as was mentioned above. For elements with still higher values of T_c/Θ, e.g., Hg (0.056) and Pb (0.075), the deviation becomes positive throughout the entire temperature range from 0°K to the transition temperature.

Calculations of these properties of superconductors which give closer agreement with the experimental data have been supplied by Swihart in a private communication. He assumed that the energy gap $\epsilon(T)/\epsilon(0)$ was the BCS function of reduced temperature and remained constant with energy out to a cut off of $13.7\epsilon(0)$ instead of to infinity as used in the weak-coupling case by BCS. The calculations were made for two cases, one assuming the ultrasonic energy gap $2\epsilon(0)/kT_c = 3.75$ and the other the value 3.84 obtained from tunneling. These cases correspond to values of T_c/Θ of 0.039 and 0.038, respectively, which are much closer to the values 0.033 and 0.038 for niobium than the weak-coupling values to which the BCS procedure is applicable.

The electronic specific heats calculated by Swihart are compared to the experimental values in Fig. 6. The values are seen to be relatively intensitive to the assumed energy gap values and both sets of calculations agree well with the experimental data for $T_c/T < 5.5$. In contrast to the specific-heat curve of BCS which is applicable only from inverse reduced temperatures greater than about 2.0, the Swihart model agrees well with experiment with T_c/T values almost to the transition temperature. For values of T_c/T greater than 5 or 6, both the BCS theory and Swihart's modification predict a downward displacement of $C_{es}/\gamma T_c$ from the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

exponential. As can be seen from Fig. 6, however, the experimental points show an upward displacement at higher inverse reduced temperatures. A similar upward departure has been observed in several other superconductors[25,26] and is generally believed to be due to anisotropy of the energy gap. Although this rise in specific heat is greater than the probable experimental error, the electronic component, C_{es}, is very small at these temperatures and approaches the order of magnitude of the scatter in the data. For this reason, no accurate quantitative comparison with the ultrasonically measured anisotropy is practicable.

Swihart's calculated critical field deviations are shown together with the experimental curve in Fig. 5. The experimental curve lies between the two calculated curves except at the low temperatures where the calculated positive deviations exceed those observed experimentally. Qualitative agreement is good as the calculations show the required change in sign, and as can be seen from the graph, the quantitative results are much closer to the experimental values than are those of the BCS theory.

For $(C_{es}/\gamma T_c)_{T_c}$ Swihart's calculations yield 2.89 for an energy gap of 3.84 and 2.76 for a gap of 3.75. Again, agreement of these values with the experimental result of 2.87 is better than the predicted BCS value of 2.43.

ACKNOWLEDGMENTS

We are greatly indebted to Professor P. H. Kessom for volunteering to carry out a specific-heat investigation on our original polycrystalline sample of Nb and to B. J. C. van der Hoeven, Jr., who collaborated in the measurements. By measurements at higher fields than available to us, they showed that below about 3°K, the value of the Debye Θ changed to 275°K. On this basis, the apparent specific-heat anomaly, similar to that in indium, is no longer present. We are indebted to Professor Serin at Rutgers University for gas thermometer determinations of the transition temperature of Pb and of a single-crystal specimen of Nb, and to Dr. James Swihart of IBM Laboratories for his generosity in sending us his specific heat and critical field calculations. Thanks are also due to Dr. Leon Schkolnick and to Bruce Biavati and Walter Schillinger for assistance with the calculations and to Dr. Robert Cotellessa, Mrs. Claire Metz, and Erwin Glück for assistance in taking and preparing the data.

[25] H. A. Boorse, Phys. Rev. Letters 2, 391 (1959).
[26] J. Bardeen and J. R. Schrieffer, *Progress in Low-Temperature Physics* (North-Holland Publishing Company, Amsterdam, 1961), Vol. 3, p. 170.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Interstitials and Vacancies in α Iron*

R. A. JOHNSON
Brookhaven National Laboratory, Upton, New York
(Received 13 January 1964)

The migration energies and atomic configurations for mono- and di-interstitials and mono- and di-vacancies in α iron have been calculated using a classical model. About 530 atoms surrounding the defect were treated as individual particles, each with three degrees of freedom, while the remainder of the crystal was treated as an elastic continuum with atoms imbedded in it. A two-body central force was devised which matched the elastic moduli, was sharply repulsive at close separation, and which went to zero midway between the second and third neighboring atoms. Configurations were found by choosing a starting configuration roughly approximating the situation under consideration and successively adjusting the value of each variable occurring in the energy equation so that the magnitude of the generalized force associated with it was zero until equilibrium was reached. The energy calculations include changes in bond energy in the discrete region, energy in the elastic field, and work done against cohesive forces, but neglect changes due to the redistribution of electrons. Calculated activation energies for motion of mono- and di-interstitials and mono- and di-vacancies were 0.33, 0.18, 0.68, and 0.66 eV, respectively, and binding energies of di-interstitials and di-vacancies were 1.08 and 0.20 eV, respectively. The stable interstitial was a "split" configuration in which two atoms were symmetrically split in a ⟨110⟩ direction about a vacant normal lattice site, and the stable di-interstitial consisted of two parallel split interstitials at nearest-neighbor lattice sites with their axes perpendicular to the line joining their centers. In the vacancy configuration an atom was missing from a normal lattice site, and the divacancy consisted of two vacancies at second-nearest-neighbor lattice sites.

INTRODUCTION

DETAILED and extensive lattice calculations for copper have been carried out for the dynamics of radiation damage events near the threshold energy for damage,[1] and for the atomic configuration[1–3] and the energy of motion[2,3] of point defects. These studies have used a similar approach to lattice calculations: The forces between atoms within a crystallite of anywhere from 50 to 1000 atoms are treated explicitly, and boundary conditions are applied to the crystallite to simulate the remainder of the lattice. The development of this method of calculation may be traced to Huntington[4] and Tewordt,[5] and a number of less extensive calculations using similar models have been performed.[6] Copper was chosen for investigation primarily because there were sufficient experimental data available, because it was felt that an appropriate and more reliable potential was available for copper than for other metals, and because the face-centered cubic structure is amenable to calculation.

The dynamic calculations[1] have recently been extended to α iron,[7] and the present research was undertaken in conjunction with that investigation. Iron was selected as the metal of greatest interest with a body-centered cubic structure. Pertinent experimental data are lacking for all body-centered cubic metals, but the threshold energy for radiation damage for α iron is known.[8]

Since the present calculation is similar to that previously used for copper by Johnson and Brown,[2] the rationale for this method will not be repeated here. In the present calculation the atoms near the defect are treated as classical particles interacting by means of a potential which applies between the first and second neighboring atoms in the lattice and which is matched to the elastic moduli.

Atomic configurations are found by choosing a starting configuration roughly approximating the situation under consideration, and then successively adjusting the value of each variable occurring in the equation for energy such that the magnitude of the generalized force associated with it is zero, and iterating this process many times. The energy in the lattice above that for a perfect lattice normally converges in the above process. Convergence is not ensured, but no difficulties have been encountered in this respect. The numerical calculations have been performed by using an IBM 7094 computer.

Iron interstitial atoms, di-interstitial pairs, vacancies, and divacancies have been studied, and the activation energy for motion as well as the atomic mechanisms of migration, as predicted by this model, have been determined for these defects. By the application of appropriate boundary conditions, the associated activation volumes for motion have also been found. Some preliminary results have been reported previously.[9]

Calculations are also being carried out for interstitial

* Work performed under the auspices of the U. S. Atomic Energy Commission.
[1] J. B. Gibson, A. N. Goland, M. Milgram, and G. H. Vineyard, Phys.Rev. 120, 1229 (1960).
[2] R. A. Johnson and E. Brown, Phys. Rev. 127, 446 (1962).
[3] A. Seeger, E. Mann, and R. v. Jan, Phys. Chem. Solids 23, 639 (1962).
[4] H. B. Huntington, Phys. Rev. 91, 1092 (1953).
[5] L. Tewordt, Phys. Rev. 109, 61 (1958).
[6] See, for example: L. A. Girifalco and V. G. Weizer, Phys. Chem. Solids 12, 260 (1960). K. H. Bennemann, Phys. Rev. 124, 669 (1961). P. Hoekstra and D. R. Behrendt, Phys. Rev. 128, 560 (1962).
[7] C. Erginsoy, G. H. Vineyard, and A. Englert, Phys. Rev. 133, A595 (1964).
[8] P. G. Lucasson and R. M. Walker, Phys. Rev. 127, 485 (1962).
[9] R. A. Johnson and A. C. Damask, Acta Met. 12, 443 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

impurities in α iron and these results will be reported elsewhere.

THEORY

Model

In the present calculation each atom within a spherical crystallite containing 531 atoms was treated as an independent particle. In the perfect lattice an atom was located at the center of the crystallite and the surrounding 530 atoms comprised 25 shells of symmetrically equivalent atoms.

The atoms outside the crystallite were treated as though they were imbedded in an elastic continuum. The displacement field $\mathbf{u}$ used in the present calculation for the elastic continuum is given by

$$\mathbf{u}=-C\nabla(1/r)=C(\mathbf{r}/r^3), \tag{1}$$

where C is the so-called "strength" of the displacement field. The displacement for a given atom i is given by

$$\mathbf{u}^i=C[\mathbf{r}_0{}^i/(r_0{}^i)^3],$$

where the subscript means that the term is to be evaluated at the perfect lattice position. Thus, C was used as the variable determining the displacement of all atoms outside the crystallite. Equation (1) gives rise to a spherically symmetric displacement field, for which the origin was taken as the center of the crystallite. The choice of the displacement field given by Eq. (1) is discussed in detail by Johnson and Brown.[2] This particular $\mathbf{u}$ is just one of an infinite number of solutions of the static, isotropic, elastic equation, but activation energies and atomic configurations near defects were found to be insensitive to other solutions. The effect of the displacement field given by Eq. (1) on the energies and configurations is relatively minor and this term could be eliminated from the calculation (i.e., the atoms outside the crystallite frozen at their perfect lattice positions) without introducing any serious changes in the results, but it was retained because it can be used to find activation volumes.

The energy of the crystallite is given by

$$E=\tfrac{1}{2}\sum_i\sum_j\varphi^{ij}+\sum_i\sum_k\varphi^{ik}+aC+bC^2, \tag{2}$$

where φ^{ij} is the interaction potential between atoms i and j, the i summation is over all atoms within the crystallite, the j summation is over all atoms within the crystallite which interact with the ith atom, and the k summation is over all atoms outside the crystallite which interact with the ith atom. The term aC accounts for the work done against the forces required to hold the perfect lattice in equilibrium, and the term bC^2 accounts for the energy stored within the elastic field. The energy for a particular configuration is given by $E-E_0$, where E_0 is the perfect lattice energy;

$$E_0=\tfrac{1}{2}\sum_i\sum_j\varphi_0{}^{ij}+\sum_i\sum_k\varphi_0{}^{ik}.$$

The derivation of the coefficients a and b will be given after the discussion of the potential.

The force on an atom i due to an atom j is given by

$$\mathbf{F}^{ij}=(\partial\varphi^{ij}/\partial r^{ij})(\mathbf{r}^{ij}/r^{ij}), \tag{3}$$

where $\mathbf{r}^{ij}=\mathbf{r}^j-\mathbf{r}^i$ (the vector from i to j). The total force on an atom i is then

$$\begin{aligned}\mathbf{F}^i&=-(\partial E/\partial\mathbf{r}^i)\\&=\sum_j\mathbf{F}^{ij}+\sum_k\mathbf{F}^{ik},\end{aligned} \tag{4}$$

where the summation indices have the same meaning as in Eq. (2). The generalized force on the crystallite arising for the elastic variable C is given by

$$\begin{aligned}F^c&=-(\partial E/\partial C)\\&=-\sum_i\sum_k\mathbf{F}^{ik}\cdot(\partial\mathbf{r}^k/\partial C)-a-2bC.\end{aligned} \tag{5}$$

The process for finding energy minima and saddle points was as follows: Initial vector positions of each atom within the crystallite were chosen to approximate the configuration of interest. Each coordinate of each atom within the crystallite was varied in turn until the corresponding force component became zero, and then the value of C was adjusted so that F^c was zero. Usually 10 to 20 such iterations were required for the energy and the configuration to converge sufficiently. The force on a given atom and the generalized elastic force are very nearly linear with displacement for small displacements. Thus, it was possible to find the force for a given variable at two values of the variable and use linear extrapolation to the value where the force is zero.

The volume expansion associated with a lattice configuration is easily calculated and is found to be linear with C, the elastic variable. Let $\Delta V'$ be the volume expansion of a hypothetical sphere around the defect.

$$\begin{aligned}\Delta V'&=\int\int\mathbf{u}\cdot d\mathbf{S}\\&=4\pi C.\end{aligned} \tag{6}$$

$\Delta V'$ is seen to be independent of the radius of the sphere. For a finite lattice, Eshelby[10] has shown that the boundary condition of zero stress at the surface of the lattice gives rise to an additional term in the volume expansion (the so-called "image force" correction). This correction may be written as

$$\Delta V=\Delta V'\{1+[4C_{44}{}^0/(3C_{12}{}^0+2C_{44}{}^0)]\}, \tag{7}$$

[10] J. D. Eshelby, J. Appl. Phys. 25, 255 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where the superscript indicates that the effective isotropic elastic constants, as given by Leibfried,[11] are to be used.

$$C_{12}^0=\tfrac{1}{5}(C_{11}+4C_{12}-2C_{44})$$
$$C_{44}^0=\tfrac{1}{5}(C_{11}-C_{12}+3C_{44}).$$

Applying these to Eq. (7) leads to

$$\Delta V=\Delta V'\left(1+\frac{4(C_{11}-C_{12})+12C_{44}}{15B}\right), \qquad (8)$$

where B, the bulk modulus, is given by $\frac{1}{3}(C_{11}+2C_{12})$.

Potential

Experimental elastic-constant data were used as a basis for obtaining an interatomic potential for the present calculation. The elastic constants of a metal may be thought of as arising primarily from two contributions: long-range electronic interactions and short range ionic interactions. The short-range interactions predominate for transition metals, but corrections were included for the long-range terms. Analytic expressions for the electronic contribution to the elastic constants of body-centered cubic metals have been derived by Fuchs[12] using the free-electron model. Since the long-range contributions are small and since the free-electron model is expected to be a rough approximation, the calculated values were used only as a guide. The elastic moduli of iron[13] and the long- and short-range contributions used in the present calculation, in units of eV/Å^3, are listed in Table I.

The simplest type of potential that can be used in a lattice calculation is one which gives rise to central forces between pairs of atoms. The ranges of the potential must not be so large as to give rise to an interaction between distant pairs or the calculation becomes too time consuming. For the body-centered cubic lattice, it is natural to cut off the potential between the second and third nearest neighbors, since the ratio of distances to the nearest, second nearest, and third nearest neighbors is $\sqrt{3}$: 2: $2\sqrt{2}$, respectively. Thus, the first and second nearest neighbor distances are comparable, while the third nearest neighbors are considerably more distant. The body-centered cubic structure is not stable with just a nearest-neighbor central force interacting between the lattice atoms, so the potential cannot be cut off between the nearest and second nearest neighbors. Since the potential is meant to describe the short-range ionic interactions, the inclusion of only nearest and second nearest-neighbor interactions is consistent with the division of the elastic-constant data into long-range and short-range contributions.

TABLE I. Elastic moduli of iron (units of eV/Å^3).

	$C_{11}-C_{12}$	C_{44}	B
Electronic contribution	0.057	0.162	0.281
Tonic contribution	0.600	0.600	0.800
Experimental values[a]	0.657	0.762	1.081

[a] See Ref. 13.

The conditions which a spherically symmetric potential $\varphi(r)$ extending through second neighbors must fulfill to match the short-range elastic moduli are

$$(C_{11}-C_{12})_{sr}=\frac{2}{3r_2}\left(\frac{6}{r_1}\varphi_1'+3\varphi_2''+\frac{3}{r_2}\varphi_2'\right), \qquad (9a)$$

$$(C_{44})_{sr}=\frac{2}{3r_2}\left(\varphi_1''+\frac{2}{r_1}\varphi_1'+\frac{3}{r_2}\varphi_2'\right), \qquad (9b)$$

$$(B)_{sr}=\frac{2}{3r_2}\left(\varphi_1''-\frac{2}{r_1}\varphi_1'+\varphi_2''-\frac{2}{r_2}\varphi_2'\right), \qquad (9c)$$

where the primes indicate differentiation with respect to r and the subscripts 1 or 2 mean that the term is to be evaluated at the nearest or second nearest-neighbor distance, respectively. There are three equations with four unknowns. Figure 1 indicates a family of possible curves of $\varphi'(r)$ which fit the conditions in Eq. (9) and which are zero midway between the second and third neighbor distances. The value and slope of $\varphi'(r)$ are determined at r_1 and r_2, and the value is determined at the cut off distance r_c, while the lines indicate how smooth curves might be fitted to these conditions. Curve II, which is roughly parabolic, is the "smoothest" of the family of curves, and no reason is seen to prefer any other to it. Indeed, it would be difficult to rationalize the choice of any of the other curves in preference to the parabolic one. Equation (9) contains

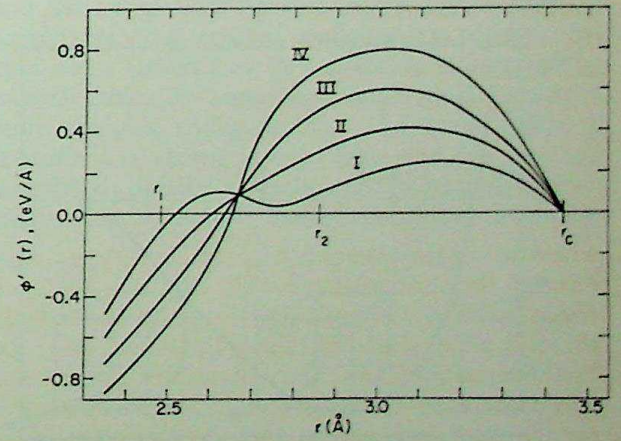


FIG. 1. A family of possible curves of the derivative of the potential, $\varphi'(r)$, which fit the short-range elastic constant conditions and which are zero midway between the second and third nearest neighbor distances.

[11] G. Leibfried, Z. Phys. 135, 23 (1953).
[12] K. Fuchs, Proc. Roy. Soc. (London) A153, 622 (1936).
[13] J. A. Rayne and B. S. Chandrasekhar, Phys. Rev. 122, 1714 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Iron potential.

Range (Å)	Potential (eV)
1.9–2.4	$-2.195976(r-3.097910)^3+2.704060r-7.436448$
2.4–3.0	$-0.639230(r-3.115829)^3+0.477871r-1.581570$
3.0–3.44	$-1.115035(r-3.066403)^3+0.466892r-1.547967$

three conditions and a parabola has three coefficients, so there is one parabola which fits the elastic moduli exactly. This parabola does not have a zero value midway between the second and third nearest-neighbor distances and so, another parabola was joined to it just past the second neighbor distance with matching value and slope, and zero value midway between the second and third neighbors. These parabolas represent $\varphi'(r)$, the derivative of the potential, and the actual potential was found by integration. The constant of integration was determined by the condition that the value of the potential be zero at the cutoff distance r_c.

Gibson *et al.*[1] and Erginsoy *et al.*[7] have shown that the threshold energy for radiation damage is determined primarily by the form of the potential at separations considerably less than the nearest-neighbor distance; distances which do not enter into the present calculation. It was felt, however, that the potential used here should be matched to the radiation damage potential for iron,[7] and thus, a cubic equation was joined with matching value and slope to the potential derived above at a distance just less than the nearest-neighbor distance, and was joined with matching value and slope to the dynamically determined potential at a distance less than any equilibrium distances expected from the calculation. Thus, the potential was a composite of the three cubic equations given in Table II and is shown in Fig. 2. All distances are given in Å, the iron lattice constant was taken as 2.86 Å, and energies are given in eV.

The coefficients a and b occurring in Eq. (2) may now be easily evaluated. From elastic theory, the pressure required to hold the crystallite in the perfect lattice configuration is[14]

$$P=\tfrac{1}{2}(C_{12}-C_{44})_{sr}$$
$$=\tfrac{1}{2}[(B)_{sr}-(C_{44})_{sr}-\tfrac{1}{3}(C_{11}-C_{12})_{sr}]. \quad (10)$$

For the elastic constants given in Table I for iron, $P=0$ and thus, $a=0$. The value of a may also be found by an explicit bond calculation in which Eq. (5) is evaluated for the perfect lattice with the requirement that $F^c=0$, i.e., that the perfect lattice is in equilibrium. Using the iron potential given in Table II, this calculation yields

$$a=0.0134 \text{ eV/Å}^3,$$

which value was used in the calculations. The discrepancy between this value of a and the value derived from elastic theory is negligible when compared with the magnitude of the elastic moduli. It arises because there is not an exact ratio of 4:3 of nearest to second nearest-neighbor bonds crossing the boundary between the crystallite and the elastic region.

The coefficient b was calculated from elastic theory[14] by setting the energy due to a displacement field $\mathbf{u}$ stored in the elastic field outside the crystallite equal to bC^2. This leads to

$$b=(64\pi^2/5Nr_2^3)[(C_{44})_{sr}+\tfrac{1}{3}(C_{11}-C_{12})_{sr}], \quad (11)$$

where N is the number of atoms in the crystallite. The pressure term aC plays a somewhat more important role than the elastic term bC^2, but for iron both terms could have been eliminated from the calculations with no serious effect upon the results.

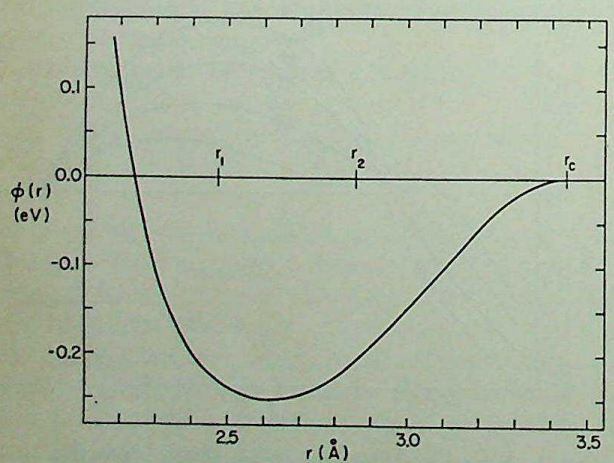


FIG. 2. The iron-iron interaction potential, $\varphi(r)$, which was used in the present calculation.

RESULTS

Interstitials

There are six interstitial configurations which must be equilibrium configurations because of symmetry, but only one of which is stable. These six configurations are denoted by the symbols I_1, $I_2 \cdots I_6$, and are described as follows: I_1 [⟨100⟩ split interstitial, Fig. 3(a)], two atoms are symmetrically split in a ⟨100⟩ direction about a vacant normal lattice site; I_2 [⟨110⟩ split interstitial, Fig. 3(b)], two atoms are symmetrically split in a ⟨110⟩ direction about a vacant normal lattice site; I_3 [⟨111⟩ split interstitial, or crowdion, Fig. 3(c)], two atoms are symmetrically split in a ⟨111⟩ direction about a vacant normal lattice site; I_4 [activated crowdion, Fig. 3(d)], an atom is located at the midpoint between two normal nearest-neighbor lattice sites; I_5 [octahedral interstitial, Fig. 3(e)], an atom is located at the midpoint between two normal second nearest-neighbor lattice sites; and I_6 [tetrahedral interstitial, Fig. 3(f)], an atom is located at the midpoint between two nearest octahedral interstitial sites.

[14] H. B. Huntington, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1958), Vol. 7.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In the three "split" cases the pairs of atoms are approximately 1.5 a_0 apart (a_0=half-lattice constant).

Configuration I_2, the ⟨110⟩ split interstitial, is the minimum energy configuration and is, therefore, stable, in agreement with the results of Erginsoy *et al.*,[7] while I_3, the ⟨111⟩ split interstitial, is just barely a local minimum, i.e., metastable. The energy and volume expansion for the various interstitial configurations are given in Table III. I_7 is the saddle-point configuration for motion of the interstitial, and is somewhat awkward to describe because it does not have much symmetry. If the initial configuration is a [110] split interstitial centered at (0,0,0), the next equilibrium configuration after one jump might be a [101] split interstitial centered at (1,1,1) a_0. The saddle point for this step is at (0.50, 0.69, 0.31) a_0. The migration sequence $I_2 \rightarrow I_7 \rightarrow I_2$ is shown in Fig. 4., and the activation energy for this process was found to be 0.33 eV. The calculation also gave an activation volume for motion of 0.1 atomic volume.

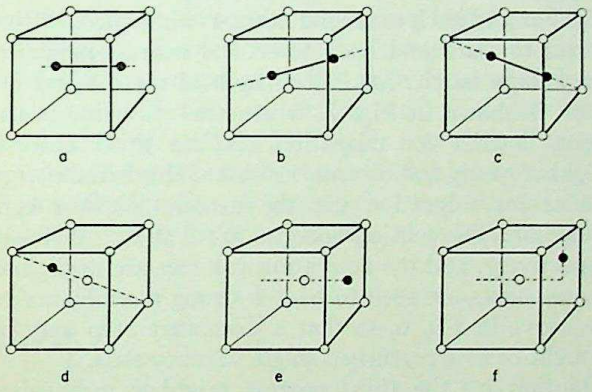


FIG. 3. Six iron interstitial configurations: (a) I_1 or ⟨100⟩ split interstitial; (b) I_2 or ⟨110⟩ split interstitials; (c) I_3, ⟨111⟩ split interstitial or crowdion; (d) I_4 or activated crowdion; (e) I_5 or octahedral interstitial; and (f) I_6 or tetrahedral interstitial.

The crowdion configuration I_3 is metastable, but by less than 0.01 eV, and the crowdion migration sequence $I_3 \rightarrow I_4 \rightarrow I_3$ has an activation energy for motion of 0.04 eV. The activation energy for rotation of a split interstitial, holding the center fixed, was found to be 0.33 eV, the same as the motion energy.

In order to make a complete study of interstitials it was necessary to check for the possible existence of other equilibrium configurations. In a manner similar to that used for copper[2] it was found to be possible to define a configuration by three coordinates. The iron calculations show that no two atoms ever approach each other closer than 1.5 a_0, and that there is always one atom within a radius of 0.75 a_0 from each normal lattice site. Thus, there is always one atom outside a bcc lattice of spheres of radius 0.75 a_0, which atom is defined as the interstitial. If the remaining atoms are at equilibrium positions consistent with the interstitial position, the crystal energy may be considered as a function of just the interstitial coordinates. If two atoms are on opposite ends of a sphere diameter, each is considered half in and half out of the sphere. These are the "split" configurations, and if one atom of a split pair enters the sphere, the other moves away from the sphere and becomes the interstitial. The volume available to the interstitial that must be studied is greatly reduced by symmetry. The volume in which the interstitial coordinates have no symmetrically equivalent position is 1/48 atomic volume, and only about half of this volume is outside the spheres, so, in effect, only about 0.01 atomic volumes must be investigated.

TABLE III. Interstitial configurations (Ω=atomic volume).

Configuration	Energy above I_2(eV)	$\Delta V(\Omega)$
I_1	1.29	1.7
I_2	...	1.6
I_3	0.32	1.7
I_4	0.36	1.7
I_5	1.12	1.4
I_6	0.85	1.5
I_7	0.33	1.7

No further equilibrium configurations of interest were found. The general pattern of the interstitial energy contours was as follows: low energy pockets at I_2 sites; a plateau region containing I_3, I_4, and I_7; and high energies elsewhere. The plateau region is roughly an oblate spheroid with its axis along the line joining two nearest neighbors. It is centered at an I_4 site, has two I_3 sites on its surface where the axis pierces the surface and has six I_7 sites on its outer rim. Not only are the configurations within this region at about the same energy, but there are no steep contours between them.

Di-Interstitials

There are many possible di-interstitial configurations and the calculations showed that many of them are metastable. Di-interstitial configurations consist of two split-single interstitials in reasonably close proximity to each other: No cases were found which resulted in more complex configurations. The most stable di-interstitial is shown in Fig. 5(a), and has two split interstitials parallel to each other at nearest-neighbor lattice sites, with their axes perpendicular to the line

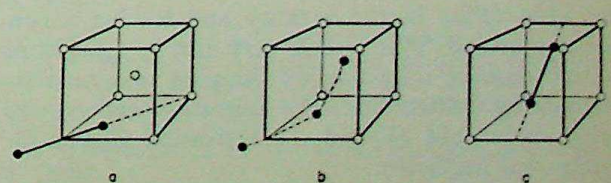


FIG. 4. The iron interstitial migration process. Configurations (a) and (c) are both I_2 configurations, and (b) is I_7, the saddle point configuration.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

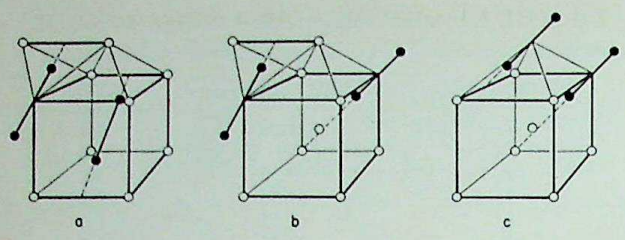


FIG. 5. The iron di-interstitial migration process. Configurations (a) and (c) are both stable di-interstitial configurations, and (b) is the intermediate step between (a) and (c).

joining their centers. The binding energy for this configuration relative to two separated interstitials was 1.08 eV and the associated volume relaxation was 0.3 atomic volumes.

The migration of di-interstitials is by a stepping process in which the two split interstitials partially dissociate. If the initial di-interstitial configuration is an interstitial split in the [110] direction at (0,0,0) a_0 and an interstitial split in the [110] direction at $(1,\bar{1},1)$ a_0, the intermediate configuration might be an interstitial split in the [110] direction at (0,0,0) a_0 and an interstitial split in the [011] direction at (2,0,2) a_0, and the final configuration would then be an interstitial split in the [011] direction at (1,1,1) a_0 and an interstitial split in the [011] direction at (2,0,2) a_0. This two step process is shown in Fig. 5, and it should be noted that the interstitial moving in each step follows a single interstitial migration path. The di-interstitial motion energy was 0.18 eV and the activation volume for motion was 0.3 atomic volumes.

The separation distance at which two split interstitials are bound as a di-interstitial did not have a well defined cutoff, but was calculated to be about 4.4 a_0. This is an average since the value depends upon the direction between the interstitials as well as their individual orientations. This range corresponds to a region of roughly 88 atomic volumes in which split interstitials have positive binding energy.

Vacancies and Di-Vacancies

The vacancy problem was straightforward compared to the interstitial calculations. The stable vacancy was the configuration in which an atom was missing from a normal lattice site, and the migration process consisted of a nearest-neighbor atom to the vacancy jumping from its normal lattice site to the vacancy site, thus filling in the vacancy and leaving a new vacancy behind. This process may also be thought of as the vacancy migrating by jumping to a nearest-neighboring lattice site. The migration energy was found to be 0.68 eV and the activation volume for motion was negligible.

The potential barrier for vacancy motion had a slight depression at the midpoint, i.e., the symmetric configuration at the midpoint of the process was not the saddle point but a local minimum. A single migration jump, therefore, had two saddle points as is shown in Fig. 6, where the energy of the configuration is plotted as a function of the position of the migrating atom. The existence of this metastable configuration should have little physical significance, however, since it is metastable by only about 0.04 eV.

A di-vacancy consists of two single vacancies in close proximity to each other. The most stable di-vacancy was that in which two vacancies were at second nearest-neighboring lattice sites, and the binding energy was 0.20 eV. The binding volume, or volume decrease of the lattice upon formation of a di-vacancy, was 0.1 atomic volumes. Vacancies at nearest neighboring sites were bound by 0.13 eV and vacancies at fourth nearest-neighboring sites were bound by 0.05 eV. No other pairs had an appreciable binding energy.

Di-vacancy migration was by a stepping process, in each step of which one of the vacancies of a di-vacancy pair moved as a single vacancy. Two possible migration processes were important; where the configuration changes from second nearest neighbor to nearest neighbor and back to second nearest neighbor, and from second to third and back to second nearest neighbor. The energy barrier for half of each of these migration paths is shown in Fig. 7, where the solid line is the energy barrier for migration via the third nearest-neighbor configuration and the dashed line is the energy barrier for migration via the nearest-neighbor configuration. The motion energies are 0.66 and 0.78 eV, respectively, and the activation volumes are negligible. These curves are seen to bear a strong resemblance to the curve in Fig. 6, so that a divacancy step may be thought of as a perturbed single vacancy step.

Motion by the third nearest neighbor mechanism always leaves the orientation of the divacancy un-

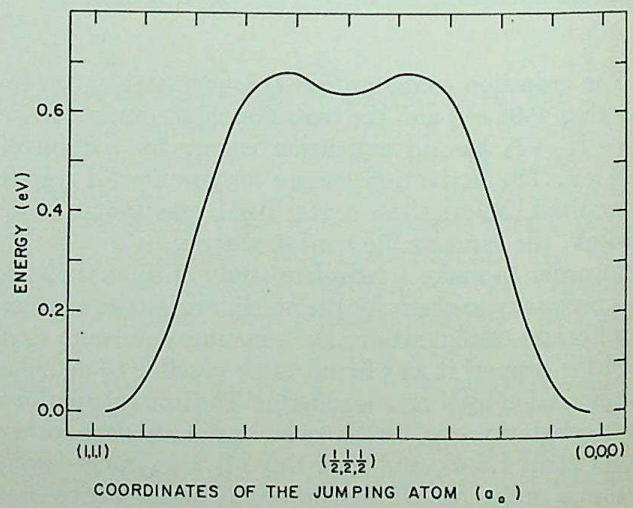


FIG. 6. The vacancy migration energy barrier. The energy of the configuration is shown as a function of the position of the jumping atom, as the vacancy migrates from (0,0,0) at the left to (1,1,1) at the right. The curve does not extend to (1,1,1) and (0,0,0) because the atom relaxes towards the vacancy.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

altered, while the orientation may be changed by a nearest neighbor process. The nearest-neighbor configuration also may act as a trap, since 0.71 eV are required for the nearest neighbor to second nearest-neighbor divacancy transition.

A summary of these results is given in Table IV. Interstitial motion is not by the same mechanism as interstitial reorientation, although the energy and volume are the same. The corresponding di-interstitial processes are the same. Vacancies have the symmetry of the lattice and therefore do not have any reorientation process associated with them.

TABLE IV. Summary of results (energy in eV, volume in atomic volumes).

	Motion energy	Motion volume	Binding energy	Binding volume
Interstitial	0.33	0.1		
Crowdion	0.04	0.0		
Di-interstitial	0.18	0.3	1.08	0.3
Vacancy	0.68	0.0		
Di-interstitial	0.66	0.0	0.20	0.1
Interstitial	0.33	0.1	Reorientation	
Di-interstitial	0.18	0.3		
Divacancy	0.78	0.0		

DISCUSSION

Results

Very little experimental data are available pertaining to the results reported in the present paper, but some comparisons may be made for interstitial migration. Lucasson and Walker[8] have obtained isochronal resistivity annealing curves for iron and copper after irradiation with electrons at energies slightly above threshold. They also found that the stage I iron and copper annealing curves can be superposed if the iron temperature scale is reduced by a factor of 2.5.[15] Corbett, Smith, and Walker[16] have found an energy value of 0.12 eV for annealing of stages I_d and I_e in copper and have ascribed this annealing to interstitial migration. Assuming that the corresponding annealing stages in iron are also due to interstitial migration, and assuming that the pre-exponential factor associated with the migration processes of the iron and copper interstitials are roughly the same, then the iron interstitial migration energy is 0.30 eV. The calculated value was 0.33 eV, in very good agreement.

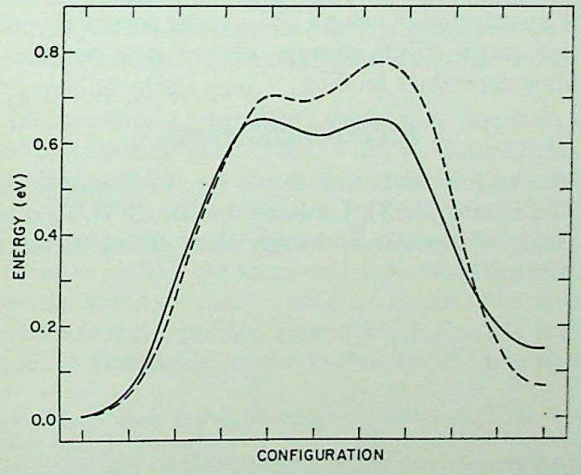


FIG. 7. The divacancy migration energy barrier. The energy of the configuration is plotted as a function of a configuration coordinate (similar to the coordinate of the jumping atom) as the divacancy transforms from a second nearest neighbor configuration on the left to a third nearest neighbor configuration along the solid line and a first nearest-neighbor configuration along the dashed line. Both curves have a mirror image repeat to the right of the drawn curves by which the migration process is completed.

In an internal fraction experiment by Wagenblast and Damask,[17] which searched for a relaxation associated with the interstitial in irradiated iron, the irradiation was performed at 140°K and the relaxation was looked for at higher temperatures, with negative results. Using the interpretation that interstitials anneal in stage I, the work of Lucasson and Walker just discussed shows that the iron interstitial anneals out of iron in ten minutes at 120°K. Thus, this interpretation is consistent with the Wagenblast and Damask findings.

Another interpretation of annealing in copper is that crowdions migrate in stages I_d and I_e and interstitials migrate in stage III.[18] The present calculations of crowdion stability and motion energy do not support this interpretation for iron. The Wagenblast and Damask experiment also does not support this interpretation since they found no indication of an oriented defect over a wide temperature range (140° to 385°K) in iron.

No direct experimental evidence is available for the motion energy of vacancies in iron. The activation energy for self-diffusion in α iron is 2.6 eV,[19] which is the sum of formation and motion energies. Theoretical estimates by Brooks[20] and by Mehl, Swanson, and Pound[21] indicate that the vacancy formation energy is considerably larger than the vacancy motion energy for bcc crystals, so that the motion energy of 0.68 eV reported in the present paper is not unreasonable compared to the self-diffusion energy. However, Damask

[15] P. G. Lucasson and R. M. Walker, Phys. Rev. 127, 1130 (1962).

[16] J. W. Corbett, R. B. Smith, and R. M. Walker, Phys. Rev. 114, 1460 (1959).

[17] H. Wagenblast and A. C. Damask, Acta Met. 10, 333 (1962).

[18] C. J. Meechan, A. Sosin, and J. A. Brinkman, Phys. Rev. 120, 411 (1960). A. Seeger, in *Proceedings of the Symposium on Radiation Damage in Solids* (International Atomic Energy Agency, Vienna, 1962), Vol. 1, p. 101.

[19] F. S. Buffington, I. D. Bakalar, and M. Cohen, J. Metals 4, 859 (1952).

[20] H. Brooks, in *Impurities and Imperfections* (American Society for Metals, Cleveland, 1955), p. 1.

[21] R. F. Mehl, M. Swanson, and G. M. Pound, Acta Met. 9, 256 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

et al.,[22] in an extensive study of carbon in α iron did not find a vacancy annealing stage at temperatures which would correspond to a motion energy of 0.68 eV, and the model which they proposed to explain their results assumes that vacancies move at a temperature corresponding to a motion energy of greater than about 1.0 eV.

The present results indicate that none of the defects studied are amenable to internal friction studies. Vacancies have the symmetry of the lattice and are, therefore, eliminated from consideration. The energy to reorient di-vacancies is greater than their motion energy, and so they will anneal before being able to reorient. Both interstitials and di-interstitials require the same energy to migrate as to reorient (migration and reorientation are the same process for di-interstitials), and thus, internal friction experiments would be very difficult to perform as the concentration of defects would be decreasing at a temperature where internal friction would be large. This does not take into account the possibility that the relaxation strength associated with these defects might be sufficiently small so that internal friction could not be measured at experimentally obtainable concentrations. A detailed calculation of this effect was made for copper[23] for an oriented interstitial with the result that even if the defect were present in heavy concentrations, it would be very difficult to detect by internal friction.

Model

The results of the present calculations are considered to accurately represent the model: Increasing the size of the region in which the atoms are allowed full freedom to relax, using greater precision within the calculation, using a different relaxation scheme, etc., would not sensibly affect the results. The important question is whether or not the model satisfactorily represents the real crystal so that the results are meaningful.

This model completely neglects any contribution from the so-called electron redistribution energy, i.e., it does not account for the drastic alteration of the electron wave functions near the defect. The author knows of no method currently available to obtain any sensible estimate of how this term enters into the differences in energy between various configurations. It is felt, however, that in all probability the calculated vacancy migration energy would be increased. It is also felt that the configurations are determined primarily by the close repulsions, so that the electron redistribution would have very little affect upon the configurations.

The energy in this model is not a sensitive function of the volume, and thus, the volume changes for the various configurations should only be considered as a rough approximation.

The choice of a potential is a critical part of any calculation, and it is felt that the potential used in the present calculation is as good an approximation as may be made. It matches the experimental elastic moduli and is fitted to a number of other not unreasonable criteria. A major objection might be that it is cut off after the second nearest neighbor. Aside from the fact that increasing the range of the potential would greatly increase the complexity of the calculation, it is desirable to test the simplest reasonable model first and only incorporate more complex features where necessary. The assumption in using a short range potential is that longer range effects average out to give rise to the volume-dependent binding.

ACKNOWLEDGMENTS

The author wishes to thank Dr. C. Erginsoy, Dr. A. C. Damask, Dr. G. J. Dienes, and Dr. G. H. Vineyard for many discussions and suggestions during the course of this work.

[22] H. Wagenblast and A. C. Damask, Phys. Chem. Solids 23, 22 (1962). F. E. Fujita and A. C. Damask, Acta Met. 12, 331 (1964). R. A. Arndt and A. C. Damask, Acta Met. 12, 341 (1964). H. Wagenblast, F. E. Fujita, and A. C. Damask, Acta Met. 12, 347 (1964).

[23] H. B. Huntington and R. A. Johnson, Acta Met. 10, 281 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Electronic Spectra of Crystalline Germanium and Silicon*†

DAVID BRUST‡
Argonne National Laboratory, Argonne, Illinois
(Received 9 December 1963)

A detailed calculation of the energy bands of germanium and silicon has been performed by use of the pseudopotential method. The first three potential coefficients have been determined empirically, and all higher ones set equal to zero. This potential was used to compute the energy eigenvalues at ~50 000 points throughout the Brillouin zone. By use of this sample, we calculated the imaginary part of the dielectric constant in the optical and near ultraviolet where direct transitions between the valence and low-lying conduction bands dominate the response. Photoelectric yield curves were obtained for comparison with recent experiments. In all cases agreement of theory and experiment was reasonable. Energy contours were constructed in several of the principal symmetry planes. These were used to explain the structure in the optical properties of Ge and Si in terms of transitions near certain important critical points. Effective masses and the static dielectric constant were also computed.

I. INTRODUCTION

RECENTLY, very precise reflectivity measurements have been taken on[1] Ge and[2] Si in the optical and near ultraviolet region of the electromagnetic spectrum. When the reflectance, $|r(\omega)|^2$, is plotted as a function of ω, the resulting curves show detailed structure. The energies involved ($1.5\ \text{eV}<\hbar\omega<10\ \text{eV}$) are such that direct interband transitions between states lying near to the forbidden gap are expected to dominate the dielectric properties of these materials.

Before attempting an explanation of the observed data, it is necessary to know the behavior of the valence and low-lying conduction bands throughout the Brillouin zone. Prior theoretical work on the energy bands of[3–5] Ge and[6–9] Si has been confined to calculations at points and along lines of particularly high symmetry. This is inadequate for our purposes; we find it necessary to extend the energy-band calculations so as to sample all of the zone.

The previous energy-band work will, however, be used as a starting point. First of all, we shall employ a simplified version of the orthogonalized plane-wave method (OPW method) utilized in Refs. 3–9. Secondly, we shall make use of the eigenvalues at symmetry points in order to deduce the numerical values of the pseudopotential matrix elements used as computational parameters in this simplified approach (see Sec. III of this paper). A reasonably precise knowledge of these eigenvalues (or rather the differences between them) is necessary to undertake the present computations.

For our purposes the energy levels can be put into two categories: levels which are sensitive to small changes in the crystal potential and those which are not. The former have, in the past, been computed *ab initio* with an uncertainty of order 3 eV,[10] while the latter were obtained to within $\sim\frac{1}{2}$ eV. The principal features (peaks and edges) of the optical curves are separated by less than 1 eV. It is necessary, therefore, to know the levels with a considerably greater accuracy than the current *a priori* calculations permit. Fortunately, a good deal of very fine experimental information is available which, when combined with the results of band theory, allows a sharp definition of both the sensitive and insensitive levels. Ehrenreich, Philipp, and Phillips[11] have given a good account of most of the main energy gaps at the symmetry points. Table I is largely taken from their work.

TABLE I. Some of the major energy gaps at the symmetry points $\Gamma=(2\pi/a)(0,0,0)$, $L=(2\pi/a)(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, and $X=(2\pi/a)(1,0,0)$. The labeling of states is after Bouckaert, Smoluchowski, and Wigner.[a] The experimental reference list is by no means exhaustive. Furthermore, energy-band calculations, not included in the table, played an important role in making the identifications.

Gap	Energy (eV) Ge	Energy (eV) Si	Refs. Ge	Refs. Si
$L_{3'}\rightarrow L_1$	2.1 2.3	3.7	1, 13, 16	17, 18, 19
$L_{3'}\rightarrow L_2$	5.9 6.1	5.5	1	2
$\Gamma_{25'}\rightarrow\Gamma_{2'}$	0.8	~3.0	17, 20	17, 21
$\Gamma_{25'}\rightarrow\Gamma_{15}$	3.1	3.5	19, 22	2, 21
$X_4\rightarrow X_1$	4.5	4.5	1	2

[a] L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

† Submitted in partial fulfillment of the requirements for the degree of Ph.D. in Physics, University of Chicago, Chicago, Illinois.

‡ Present address: Department of Physics, Purdue University, Lafayette, Indiana.

[1] H. R. Philipp and E. A. Taft, Phys. Rev. 113, 1002 (1959).
[2] H. R. Philipp and E. A. Taft, Phys. Rev. 120, 37 (1960).
[3] F. Herman and J. Callaway, Phys. Rev. 89, 518 (1952).
[4] F. Herman, Physica 20, 801 (1954).
[5] F. Herman, Phys. Rev. 93, 1214 (1954).
[6] T. Woodruff, Phys. Rev. 103, 1159 (1956).
[7] F. Bassani, Phys. Rev. 108, 263 (1957).
[8] F. Bassani, Nuovo Cimento 13, 244 (1959).
[9] L. Kleinman and J. C. Phillips, Phys. Rev. 118, 1152 (1960).
[10] F. Bassani and M. Yoshimine, Phys. Rev. 130, 20 (1963).
[11] H. Ehrenreich, H. R. Philipp, and J. C. Phillips, Phys. Rev. Letters 8, 59 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

It should be emphasized at this point that Ehrenreich *et al.* have deduced some of the energy gaps of Table I with the aid of several natural assumptions. They assume, for example, that the main features of the optical curves can be associated with direct transitions between levels at symmetry points. One of the primary purposes of the present work is to attempt a justification of such assumptions and working hypotheses. In order to illustrate further, Ehrenreich *et al.*[11] make use of certain characteristic properties of the reflectivity, as discussed by Phillips,[12] in order to make their assignments in a definite way. For example, the edge near 2.2 eV in Ge has been resolved into a doublet[13] (see Table I). This corresponds to the expected 0.2 eV spin-orbit splitting of the $L_{3'}$ level[14] provided the optical edge is taken as arising from transitions near $L_{3'} \rightarrow L_1$. This assignment was considered as reasonable since both band theory and g-factor information[15,16] indicate an $L_{3'} \rightarrow L_1$ splitting ~ 2 eV.

With the $L_{3'} \rightarrow L_1$ gap determined in Ge it was possible to deduce it in Si.[12] This was done by extrapolating the indirect gap ($\Gamma_{25'} \rightarrow L_1$) in Ge-rich Ge-Si alloys[17,18] to its value in pure Si, and also by extrapolating the $L_{3'} \rightarrow L_1$ optical edge[19] to give a value for the $L_{3'} \rightarrow L_1$ splitting in Si.

The $\Gamma_{25'} \rightarrow \Gamma_{2'}$ splitting was found in Ge by observing the threshold for direct optical transitions.[17,20] Its value could be deduced in Si by the extrapolation of alloy data[17] and also (see Ref. 12) by cyclotron resonance measurements in strained p-type Si.[21] The last-mentioned experiments as well as optical reflectivity were used to find a value for $\Gamma_{25'} \rightarrow \Gamma_{15'}$ in Si. Again cyclotron resonance[22] and extrapolation techniques[19] were used to get $\Gamma_{25'} \rightarrow \Gamma_{15}$ in Ge.[11,12] Finally, $X_4 \rightarrow X_1$ and $L_{3'} \rightarrow L_3$ were determined by comparing band-theory results (these two gaps are both insensitive to the potential) with reflectivity data.

To sum up, the energy levels of Table I provide a model on the basis of which we can begin our detailed calculations. Hopefully we will be able to explain the frequency dependence of the optical properties. This will enable us to establish the validity of Table I and provide us with a much broader understanding of the band structure than has heretofore been possible. Furthermore, the present work in the course of checking the level assignments of Ehrenreich *et al.*[11] and Phillips[12] will confirm or redefine their diagnostic techniques.

Spin-orbit effects will not be accounted for in this paper. They are observable in Ge, as mentioned, but are $<10\%$ of the band separations. These effects, if desired, can be treated as a perturbation, and their influence on the line shape determined.

II. OPTICAL PROPERTIES

A. General Discussion

If $\epsilon_2(\omega)$ is used to denote the imaginary part of the complex dielectric constant, we have from Eq. (9) of Ref. 23

$$\epsilon_2(\omega)=\frac{4\pi^2e^2\hbar}{3m^2\omega^2}\sum_{n,s}\int_{\text{B.Z.}}\frac{2}{(2\pi)^3}\times\delta(\omega_{n,s}(\mathbf{k})-\omega)\,|M_{n,s}(\mathbf{k})|^2d^3k. \quad (1)$$

The subscripts n and s refer to filled and unfilled bands, respectively, and $\omega_{n,s}(\mathbf{k})=(E_s(\mathbf{k})-E_n(\mathbf{k}))/\hbar$. $|M_{n,s}(\mathbf{k})|^2=|(U_{\mathbf{k},n}|\nabla|U_{\mathbf{k},s})|^2$, where $U_{\mathbf{k},n}$ and $U_{\mathbf{k},s}$ represent the periodic parts of Bloch functions. Expression (1) neglects lifetime broadening effects such as those resulting from phonon and impurity scattering. If we ignore for a moment the matrix element in the integral (and the factor $4\pi^2e^2\hbar/3m^2\omega^2$) then the contribution from a given pair of bands to $\epsilon_2(\omega)$ is simply

$$J_{n,s}(\omega)=\int_{\text{B.Z.}}\frac{2}{(2\pi)^3}\delta(\omega_{n,s}(\mathbf{k})-\omega)d^3k. \quad (2)$$

The quantity $J_{n,s}(\omega)$ is the joint density of states for the two bands indexed by n and s. $J_{n,s}(\omega)\Delta\omega$ is equal to the number of pairs of states in bands n and s with $\hbar(\omega-\Delta\omega/2)\leqslant(E_s(\mathbf{k})-E_n(\mathbf{k}))<\hbar(\omega+\Delta\omega/2)$. Later in our discussion of momentum matrix elements we shall see that $|M_{n,s}(\mathbf{k})|^2$ can, to a good approximation, be treated as a constant. This implies that a knowledge of the relevant $J_{n,s}(\omega)$ is all that one requires to understand the properties of $\epsilon_2(\omega)$.

In the Introduction it was stated that the reflectivity data were used as an aid in establishing the model term scheme of Table I. Actually the reflectance is a function of both $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$. However, by examining the figures of Refs. 1 and 2 and comparing them with $\epsilon_2(\omega)$ for Ge and Si presented in Sec. VIII of this paper one sees that $R(\omega)\sim\epsilon_2(\omega)$. All of the principal features in $|r(\omega)|^2$ are reproduced at the same energies in $\epsilon_2(\omega)$.

B. Theory of Optical Structure

At this point it is clear that to understand $\epsilon_2(\omega)$ (we shall direct our attention to the imaginary part of the dielectric constant hereafter), we must carefully examine the functions $J_{n,s}(\omega)$. Expression (2) can be

[12] J. C. Phillips, Phys. Rev. **125**, 1931 (1962).
[13] J. Tauc and A. Abraham, Phys. Chem. Solids **20**, 190 (1961).
[14] L. Liu, Phys. Rev. **126**, 1317 (1962).
[15] L. Roth and B. Lax, Phys. Rev. Letters **3**, 217 (1959).
[16] G. Feher, D. K. Wilson, and E. A. Gere, Phys. Rev. Letters **3**, 25 (1959).
[17] R. Braunstein, A. R. Moore, and F. Herman, Phys. Rev. **109**, 695 (1958).
[18] E. R. Johnson and S. M. Christian, Phys. Rev. **95**, 560 (1954).
[19] J. Tauc and A. Abraham, Phys. Rev. Letters **5**, 253 (1960).
[20] S. Zwerdling and B. Lax, Phys. Rev. **106**, 51 (1956).
[21] J. C. Hensel and G. Feher, Phys. Rev. Letters **5**, 307 (1960).
[22] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368. (1955).
[23] M. H. Cohen, Phil. Mag. **3**, 762 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

transformed to

$$J_{n,s}(\omega)=\frac{2}{(2\pi)^3}\int_{\omega_{n,s}(\mathbf{k})=\omega} ds/|\nabla_k\omega_{n,s}(\mathbf{k})|\,, \qquad (3)$$

where the integral extends over the surface defined by $\omega_{n,s}(\mathbf{k})=\omega$ and ds is an infinitesimal element of area on that surface. The points $\mathbf{k}_0$ where $|\nabla_k\omega_{n,s}(\mathbf{k})|_{\mathbf{k}=\mathbf{k}_0}=0$ are significant and are called critical points (c.p.). Van Hove[24] wrote the following expansion, valid about a normal c.p.,

$$\omega_{n,s}(\mathbf{k})=\omega_c+\sum_{\alpha=1}^{3} a_\alpha\epsilon_\alpha\Delta\mathbf{k}_\alpha^2\,, \qquad (4)$$

where $\epsilon_\alpha=\pm1$, $\Delta\mathbf{k}=\mathbf{k}-\mathbf{k}_0$. Here $\omega_c=\omega_{n,s}(\mathbf{k}_0)$, and a coordinate transformation which preserves d^3k has been made so as to get the functional dependence expressed by (4). He found that the critical points produce analytic singularities in what we call $J_{n,s}(\omega)$. The mathematical behavior of $J_{n,s}(\omega)$ near the "Van Hove" singularities is described by the following list:

M_0 $\quad\epsilon_1=\epsilon_2=\epsilon_3=1\,,$

$$J_{n,s}(\omega)=\begin{cases}C+O(\omega-\omega_c) & \text{when } \omega<\omega_c\\ C+A(\omega-\omega_c)^{1/2}+O(\omega-\omega_c) & \text{when } \omega>\omega_c\,,\end{cases} \qquad (5)$$

M_1 $\quad\epsilon_1=\epsilon_2=-\epsilon_3=1\,,$

$$J_{n,s}(\omega)=\begin{cases}C-A(\omega_c-\omega)^{1/2}+O(\omega_c-\omega) & \text{when } \omega<\omega_c\\ C+O(\omega-\omega_c) & \text{when } \omega>\omega_c\,,\end{cases} \qquad (6)$$

M_2 $\quad\epsilon_1=\epsilon_2=-\epsilon_3=-1\,,$

$$J_{n,s}(\omega)=\begin{cases}C+O(\omega-\omega_c) & \text{when } \omega<\omega_c\\ C-A(\omega-\omega_c)^{1/2}+O(\omega-\omega_c) & \text{when } \omega>\omega_c\,,\end{cases} \qquad (7)$$

M_3 $\quad\epsilon_1=\epsilon_2=\epsilon_3=-1\,,$

$$J_{n,s}(\omega)=\begin{cases}C+A(\omega_c-\omega)^{1/2}+O(\omega-\omega_c) & \text{when } \omega<\omega_c\\ C+O(\omega-\omega_c) & \text{when } \omega>\omega_c\,.\end{cases} \qquad (8)$$

In each of these expressions C and A represent constants. The symbol M_0 designates a minimum in $\omega_{n,s}(\mathbf{k})$, M_3 denotes a maximum, and M_1 and M_2 refer to saddle points. In Figs. 1(a)–1(d) we plot the behavior of $J_{n,s}(\omega)$ near the Van Hove singularities. One can already see the way in which a critical point can produce *edges* in the joint density of states and hence in $\epsilon_2(\omega)$. In Fig. 1(e) we show the possibility of two nearly degenerate critical points of the M_1 and M_2 type producing a peak-like structure in $J_{n,s}(\omega)$. Later work by Phillips[25] showed that the periodicity of the energy bands requires the existence of a minimal set of critical points. From group theoretical arguments he demonstrated that

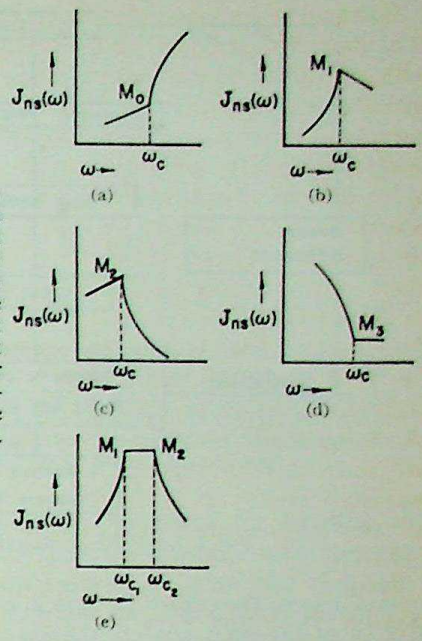


FIG. 1. Joint density of states near a normal critical point (a) M_0, (b) M_1, (c) M_2, (d) M_3, (e) a nearly degenerate pair M_1 and M_2. It should be noted that the slope of the linear portion is not determined by the nature of the critical point.

critical points occur at symmetry points. In the simplest cases (e.g., lattice vibration spectra), almost all the critical points may occur at symmetry points. In the absence of detailed information about $E(\mathbf{k})$ throughout the Brillouin zone, the empirical analysis of the previous section was confined to symmetry points. For the complicated band structures of Ge and Si this simplification (which has also been made in order to interpret the ultraviolet spectra of the noble metals) is certainly not valid. Moreover, the constants in (5)–(8) which indicate the strength of a given edge can only be obtained from values of $E(\mathbf{k})$ throughout the Brillouin zone.

C. Band-Structure Approach to Optical Properties

We now wish to do a band-structure calculation with the view of deriving $\epsilon_2(\omega)$. A description of the logical framework to be employed is the following. First a method of computing the bands will be proposed. The method should have a small number of disposable parameters. In order to determine these parameters, the term scheme discussed in Sec. I will be used as a starting point. Having set the parameters, we shall then proceed to find the bands throughout the Brillouin zone. From the band structure $\epsilon_2(\omega)$ can be computed. If the dielectric function agrees with experiment, we shall conclude that both the model of Sec. I and the computed band structure are correct. Should there be substantial disagreement then a consistent alternate to the starting term scheme will be tried, and a corrected band structure found. A schematic diagram of this is shown in Fig. 2.

Since the optical structure is associated with critical points, we must find them in the computed bands. Of course, we know that they will appear at the symmetry points, and they will have energies which correspond to

[24] L. Van Hove, Phys. Rev. 89, 1189 (1953).
[25] J. C. Phillips, Phys. Rev. 104, 1263 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

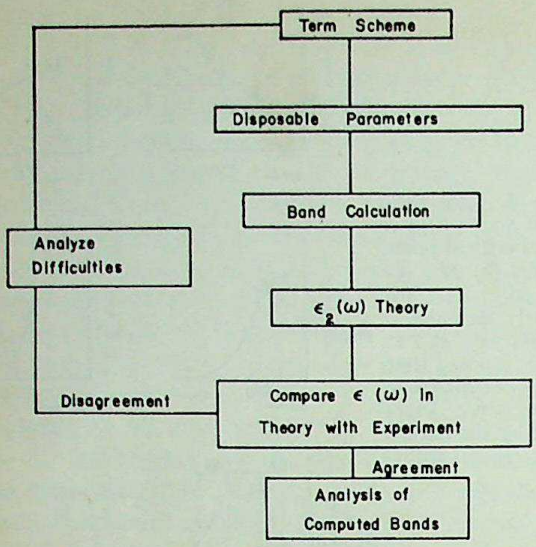


FIG. 2. Logical flow diagram.

optical edges. We cannot be sure in advance that they will be of the right analytical class or will contribute edges of the correct strength. Furthermore, there may be critical points away from the principal symmetry points, in which case they will have to be located.

III. PSEUDOPOTENTIAL METHOD

We mentioned earlier that for certain levels the OPW technique is only accurate to ~ 3 eV.[10] Furthermore, the unwieldy nature of the orthogonalization terms makes it unsuitable for extensive calculations. These considerations imply that the OPW method is unsuited for our needs.

Research workers using the OPW method noticed that the effect of orthogonalizing the plane waves to core orbitals was to greatly diminish the magnitude of the off-diagonal matrix elements in the secular equation. This led Phillips and Kleinman[26] to write the wave functions in the form

$$\psi_{n,\mathbf{k}}(\mathbf{r})=\phi_{n,\mathbf{k}}(\mathbf{r})-\sum_l \langle\theta_{l,\mathbf{k}}(\mathbf{r})|\phi_{n,\mathbf{k}}(\mathbf{r})\rangle\theta_{l,\mathbf{k}}(\mathbf{r}). \quad (9)$$

Here $\theta_{l,\mathbf{k}}$ is the core function constructed from the lth core orbital, and $\phi_{n,\mathbf{k}}$ is a smooth wave function satisfying the wave equation

$$(P^2/2m+V_c+V_R)\phi_{n,\mathbf{k}}=E_n(\mathbf{k})\phi_{n,\mathbf{k}}, \quad (10)$$

where V_c is the crystal potential. The operator V_R has the character of a repulsive potential. Cohen and Heine[27] show that $\phi_{n,\mathbf{k}}$ and V_R are not unique. By a variational argument they demonstrated that the latter can be chosen so as to minimize

$$V_{\text{eff}}=V_c+V_R, \quad (11)$$

in which case V_R has the form

$$V_R\phi_{n,\mathbf{k}}=-\sum_l(\theta_{l,\mathbf{k}}|V_c\phi_{n,\mathbf{k}})\theta_{l,\mathbf{k}}. \quad (12)$$

[26] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959).
[27] M. H. Cohen and V. Heine, Phys. Rev. **122**, 1821 (1961); also, see Ref. 30.

From (11) it is apparent that if the set of core orbitals were complete, perfect cancellation of $(V_c\phi_{n,\mathbf{k}})$ by $(V_R\phi_{n,\mathbf{k}})$ would result (the bands would then be completely free electron). The core orbitals form a surprisingly good basis set in the core region. This implies that the Fourier coefficients of $V_{\text{eff}}^{\mathbf{G}_i}=V_c^{\mathbf{G}_i}+V_R^{\mathbf{G}_i}$ are small for large values of $\mathbf{G}_i=(a/2\pi)\mathbf{K}_i$ where $\mathbf{K}_i$ is a reciprocal lattice vector.[28] This suggests introducing a pseudopotential with Fourier coefficients $V_p^{\mathbf{G}_i}\simeq V_{\text{eff}}^{\mathbf{G}_i}$ which are zero for $|\mathbf{G}_i|$ greater than the first two or three values.[29] A pseudopotential of this form was first introduced by Phillips.[30] He, however, chose $V_p^{\mathbf{G}_i}$ = const for large $\mathbf{G}_i$. Bassani and Celli,[31] on the other hand, took $V_p^{\mathbf{G}_i}=0$ for $\mathbf{G}_i$ such that $|\mathbf{G}_i|^2>11$ which materially improved the convergence in agreement with the variational arguments of Cohen and Heine.

In this paper we shall adopt a pseudopotential having the form used by Bassani and Celli. For both Ge and Si this will be of the form

$$V_p=\sum_j V_p^{\mathbf{G}_j}\exp[(2\pi i/a)\mathbf{G}_j\cdot\mathbf{r}],$$

$$V_p^{\mathbf{G}_j}=V_p(3)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=3,$$

$$=V_p(8)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=8,$$

$$=V_p(11)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=11,$$

$$=0 \quad \text{if} \quad |\mathbf{G}_j|^2>11.$$

Here $\mathbf{G}_{j,1}$ is the projection of $\mathbf{G}_j$ on the (1,0,0) direction (direction normal to a square face). There are then three adjustable parameters $V_p(3)$, $V_p(8)$, and $V_p(11)$. We should point out that V_{eff} is both l and $\mathbf{k}$ dependent. That is, it is different depending on whether $\phi_{n,\mathbf{k}}$ has s or p symmetry, and also varies throughout the Brillouin zone. This nonlocal character is a consequence of the OPW method. We notice, however, that V_p is a local (function of $\mathbf{r}$ only) potential. This simplifying feature cannot be regarded as deleterious. The energy eigenvalues which result from using V_p are generally within 2 eV of those deriving from the use of V_{eff}. The latter, one remembers, gives energies certain to only 3 eV furthermore, V_p will be adjusted at the outset to agree with what we believe to be a large number of symmetry term values. We would then expect V_p to give eigenvalues at other points of the zone with an error much

[28] In these units the center of the hexagonal face is located at $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and the shortest nonzero $\mathbf{G}_i$ has length $=\sqrt{3}$.
[29] We are here distinguishing between V_{eff} which is uniquely defined by (7) and (8) and V_p which is an approximation to V_{eff}. Inserting V_{eff} into (6) gives exactly the same results as the OPW method from which it is derived, whereas V_p does not.
[30] J. C. Phillips, Phys. Rev. **112**, 685 (1958).
[31] F. Bassani and V. Celli, Phys. Chem. Solids **20**, 64 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

less than the OPW method. The pseudopotential in a sense circumvents the task of finding a crystal potential replete with a correct exchange interaction, Coulomb interaction, etc. It fits, instead, the most important terms in the potential to experiment.

One more point should be made. Bassani and Celli do distinguish between s- and p-like states by allowing $V_s(0)$ to differ from $V_p(0)$. In this way they can raise s-like states with respect to p-like states. We have found that this is not necessary to get an adequate fit to the experimental interpretation of the levels.

IV. DIAGONALIZATION PROCEDURE

We seek to determine the energy eigenvalues of our model wave equation

$$H_p\phi_{n,\mathbf{k}}=E_{n,\mathbf{k}}\phi_{n,\mathbf{k}}, \tag{13}$$

where $H_p=(-\hbar^2/2m)\nabla^2+V_p$. To do this we must in principle solve the infinite secular equation

$$|H_p{}^{i,j}-E\delta_{i,j}|=0. \tag{14}$$

Here $H_p{}^{i,j}=\langle\phi_{\mathbf{k}+\mathbf{K}_i}|H_p|\phi_{\mathbf{k}+\mathbf{K}_j}\rangle$ and $\phi_{\mathbf{k}+\mathbf{K}_i}=e^{i(\mathbf{k}+\mathbf{K}_i)\cdot\mathbf{r}}$. That is, we use a plane-wave representation in which to expand the $\phi_{n,\mathbf{k}}$. The lowest four levels will represent valence states and the higher one's conduction states (one remembers that the core levels have been eliminated from the problem by the orthogonalization procedure). Only the low-lying conduction levels are of interest so that we shall ask for only the first eight eigenvalues of H_p (four valence and four conduction).

Assuming that the $\phi_{\mathbf{k}+\mathbf{K}_i}$ have been ordered so that for $i>j$, $|\mathbf{k}+\mathbf{K}_i|^2\geq|\mathbf{k}+\mathbf{K}_j|^2$, we may truncate the secular Eq. (14) so as to include contributions from only those plane waves for which $i,\ j\leq N$ (N some integer). We expect the low-lying levels to converge to their final values as N is made large. If we take as our convergence criterion that $|E_{n,\mathbf{k}}{}^{N=N_0}-E_{n,\mathbf{k}}{}^{N=\infty}|<0.1$ eV then N_0 would have to be $\sim$50. The time required by the IBM 704 to diagonalize a 50×50 matrix at a large number of points was judged as too great to make this approach feasible. By augmenting the calculation with perturbation theory it is possible, however, to reduce greatly the computing time.

A form of perturbation theory, e.g., Löwdin[32] was used instead of the Rayleigh-Schrödinger method. We seek the eigenvalues of the secular equation

$$|U_p{}^{n,m}-E\delta_{n,m}|=0, \tag{15}$$

$$U_p{}^{n,m}=H_p{}^{n,m}+\sum_{\gamma=N+1}^{\Gamma}\frac{H_p{}^{n,\gamma}H_p{}^{\gamma,m}}{E-H_p{}^{\gamma,\gamma}}, \tag{16}$$

with $n,\ m\leq N$ and $\gamma>N$ but $\leq\Gamma$ where N and Γ are integers. The indexes n and m refer to plane waves being treated exactly; whereas Γ refers to higher plane waves being treated only through perturbation theory. This procedure is especially convenient for handling degenerate and quasidegenerate cases as it is not necessary to find linear combinations of zero-order degenerate states as in the standard method. This is particularly advantageous when automatic computing machinery is employed.

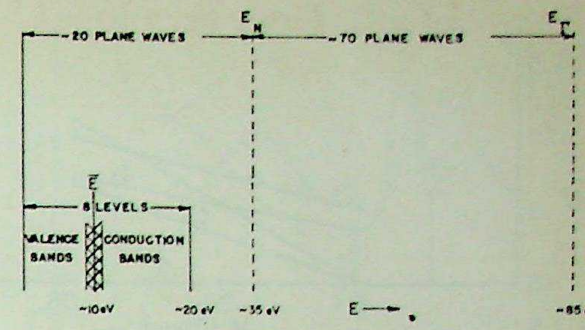


FIG. 3. The approximate relationship of the quantities entering the band calculation. $\bar{E}$ represents an average energy for the valence and conduction bands. E_N is a cutoff energy such that all plane waves with kinetic energy $(\hbar^2/2m)|\mathbf{k}+\mathbf{K}_i|^2\leq E_N$ are treated exactly, whereas those having a kinetic energy between E_N and E_Γ are accounted for by perturbation theory.

Before proceeding it was necessary to eliminate the eigenvalue dependence of the matrix elements in (15). If this were not done, we should have to solve for each of the eigenvalues separately and to use an iteration procedure for each of them. We wished, instead, to diagonalize (15), and obtain all eight interesting levels at once.

In handling the matrix elements the following substitutions were made: off the diagonal $\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,m})/(E-H_p{}^{\gamma,\gamma})\rightarrow\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,m})/(E-H_p{}^{\gamma,\gamma})$, that is, $E\rightarrow\bar{E}$. Here $\bar{E}$ is an average of the eight energy levels at each point in the reduced zone. On the diagonal $\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,n})/(E-H_p{}^{\gamma,\gamma})\rightarrow\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,n})/(H_p{}^{n,n}-H_p{}^{\gamma,\gamma})$, i.e., $E\rightarrow H_p{}^{n,n}=(\hbar^2/2m)|\mathbf{k}+\mathbf{K}_n|^2$.

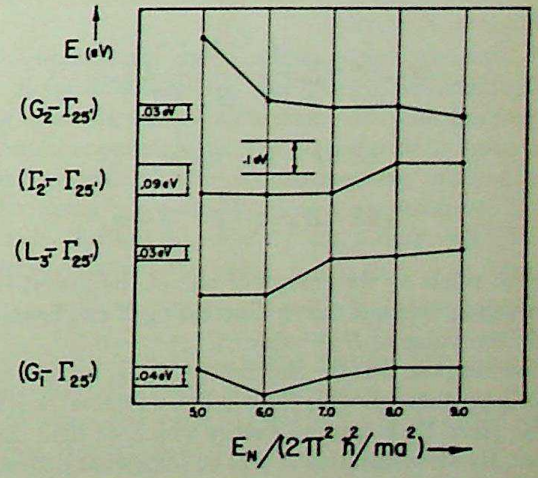


FIG. 4. Shows the convergence properties of the worst cases among those tested. G_1 and G_2 refer to states at the general point of the zone with $\mathbf{k}=(0.60,0.35,0.10)$. The energies (in eV) at the left are an estimate of the convergence for $E_N=7.0$ (the value actually used). They were gotten by comparing with the results for $E_N=9.0$.

[32] P. Löwdin, J. Chem. Phys. 19, 1396 (1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

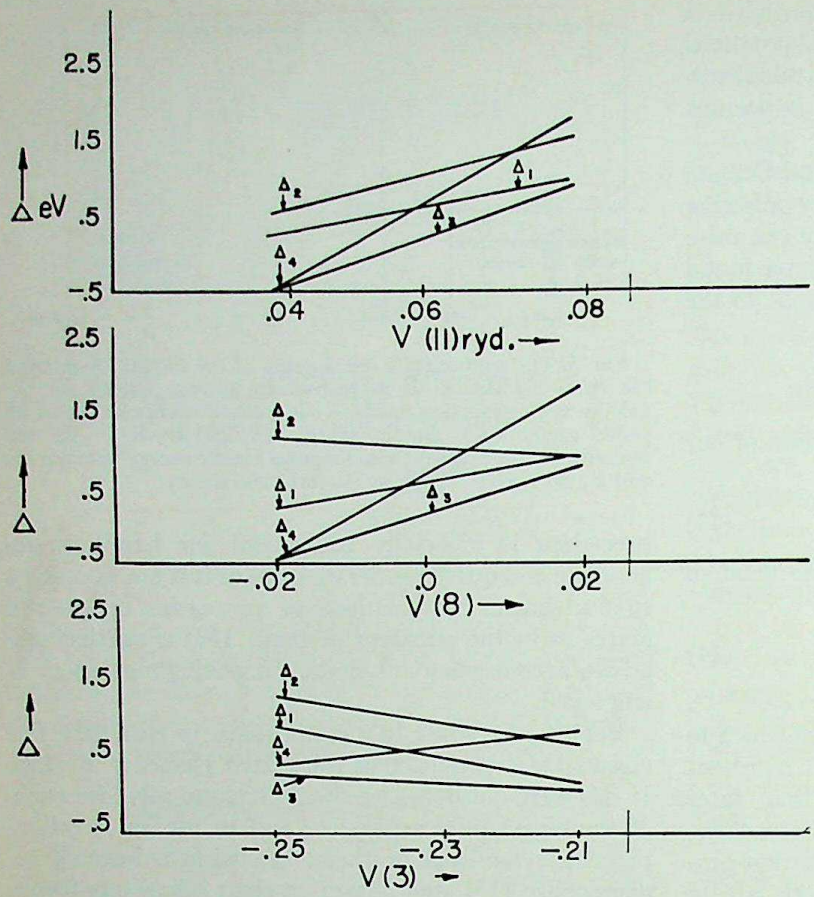


FIG. 5. Variation of some of the principal band gaps with the potential coefficients. In eV: $\Delta_1=(X_1-X_4)-3.0$, $\Delta_2=(L_3-L_{3'})-4.5$, $\Delta_3=(L_1-L_{3'})-1.5$, $\Delta_4=(\Gamma_{2'}-\Gamma_{25'})$; (a) $V(3)=-0.21$ Ry, $V(8)=0.0$ Ry; (b) $V(3)=-0.23$ Ry, $V(11)=0.06$ Ry; (c) $V(8)=0.0$ Ry, $V(11)=0.06$ Ry.

We here, in essence, replace E by the kinetic energy of the principal plane waves making up the expansion of $\phi_{n,\mathbf{k}}$. This allows a closer representation of E than any one choice of $\bar{E}$ permits. It was felt that the more accurate representation of the diagonal matrix elements would improve the convergence.

The error $\Delta U_p{}^{n,m}$ in a matrix element resulting from these approximations is

$$\Delta U_p{}^{n,m}\approx\sum_\gamma \frac{H_p{}^{n,\gamma}H_p{}^{\gamma,m}}{(E-H_{\gamma,\gamma}{}^p)^2}\delta E,$$
$$\delta E=\bar{E}-E, \quad n\neq m, \tag{17}$$
$$\delta E=H_p{}^{n,m}-E, \quad n=m.$$

Since E refers to the energy of one of the lowest eight eigenvalues, one sees that by increasing N and hence the minimum value of $H_p{}^{\gamma,\gamma}$ that $\Delta U_p{}^{n,m}\to 0$.

Now suppose that the plane waves $\phi_{\mathbf{k}+\mathbf{K}_n}$ being treated exactly are those obeying the condition $(\hbar^2/2m)|\mathbf{k}+\mathbf{K}_n|^2 \leqslant E_N$. Here E_N is a cutoff energy which we shall choose below. By this definition N will be somewhat dependent on $\mathbf{k}$. In a like manner the plane waves $\phi_{\mathbf{k}+\mathbf{K}_\gamma}$ entering the calculation through perturbation theory satisfy $E_N<(\hbar^2/2m)|\mathbf{k}+\mathbf{K}_\gamma|^2\leqslant E_\Gamma$.

The actual choice of E_N, E_Γ, and $\bar{E}$ depend on an intelligent use of trial and error. The values ultimately chosen were

$$E_N=7.0(2\pi^2\hbar^2/ma^2)\sim 35 \text{ eV},$$
$$E_\Gamma=19.0(2\pi^2\hbar^2/ma^2)\sim 85 \text{ eV},$$
$$\bar{E}=2.0(2\pi^2\hbar^2/ma^2)\sim 10 \text{ eV}.$$

With this choice N was $\sim$20 and $\Gamma\sim$90. In Fig. 3 we show these values in comparison with the band energies.

The question of convergence was examined by observing the change in the eigenvalues as E_N was increased. In Fig. 4 the levels showing the poorest convergence among the test cases are plotted as a function of E_N. One should note that convergence is good to $\sim$0.05 eV which is within the limits of the accuracy being sought.

V. PSEUDOPOTENTIAL PARAMETERS

We now turn to a brief discussion concerning the choice of the pseudopotential coefficients. The work of Bassani and Celli[31] on Ge was found to be a convenient starting point. Since the present work makes no distinction between s and p symmetry, it was necessary to redefine their potential slightly. In determining the Fourier coefficients particular emphasis was placed on getting good agreement with the proposed interpretation of the optical experiments. That means we sought to make the direct splittings such as $L_{3'}\to L_3$, $L_{3'}\to L_1$, etc., come as close as possible to the experimental

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Energy of principal gaps (in eV).

Gap	Bassani (Ge)	Brust (Ge)	Brust (Si)
$\Gamma_{25'} \rightarrow \Gamma_{2'}$	0.8	0.6	3.8
$\Gamma_{25'} \rightarrow \Gamma_{15}$	3.4	3.6	3.4
$L_{3'} \rightarrow L_1$	2.0	1.8	3.1
$L_{3'} \rightarrow L_3$	5.4	5.4	5.4
$X_4 \rightarrow X_1$	4.4	3.6	4.0

identification. Since at our first trial we sought to test the proposed energy model by attempting to get the correct structure in the optical properties, it was natural to make the splittings agree as closely as possible with the model.

In order to reduce the labor required to get a good set of pseudopotential parameters the graph shown in Fig. 5 was prepared. This indicates how the principal band gaps vary with changes in the potential coefficients. By using this a good fit could be obtained with only one or two trials for both Ge and Si. In this way we found for Ge: $V(3)=-0.23$ Ry, $V(8)=0.0$ Ry, and $V(11)=0.06$ Ry; while for Si: $V(3)=-0.21$ Ry, $V(8)=0.04$ Ry, and $V(11)=0.08$ Ry. The term values resulting from these choices are given in Table II where our Ge results are compared with those of Bassani and Celli.

The values of Table II are somewhat different from those of Table I. It is not possible to get an exact fit with a three-parameter potential. If one required a better set of values, additional coefficients would have to be introduced into the potential.

It was of interest to compute the energy bands along symmetry directions. These are shown in Fig. 6, and should be compared with earlier results obtained by Herman on Ge and by Kleinman and Phillips for Si. We should point out that the good behavior of our bands was used as an indication of both the convergence of our computational procedure and of the correctness of the pseudopotential approach.

We are now in a position to go on to a calculation of $\epsilon_2(\omega)$. Before doing this, however, we shall examine the result of computing effective masses according to the present scheme.

VI. EFFECTIVE MASSES

We wish to determine the effective masses of the conduction electrons within the framework of the pseudopotential method. Doing this for Ge by the formulas of $\mathbf{k}\cdot\mathbf{p}$ perturbation theory we have for Ge, e.g., Liu[14]

$$\left(\frac{m_e}{m_t^*}\right)=1+\frac{2}{m}\frac{|\langle L_1|P_{y'}|L_{3'}{}^{y'}\rangle|^2}{(E_{L_1}-E_{L_{3'}})}. \qquad (18)$$

Here $P_{y'}$ is a momentum operator in a direction transverse to the (1,1,1) direction. Using our pseudopotential

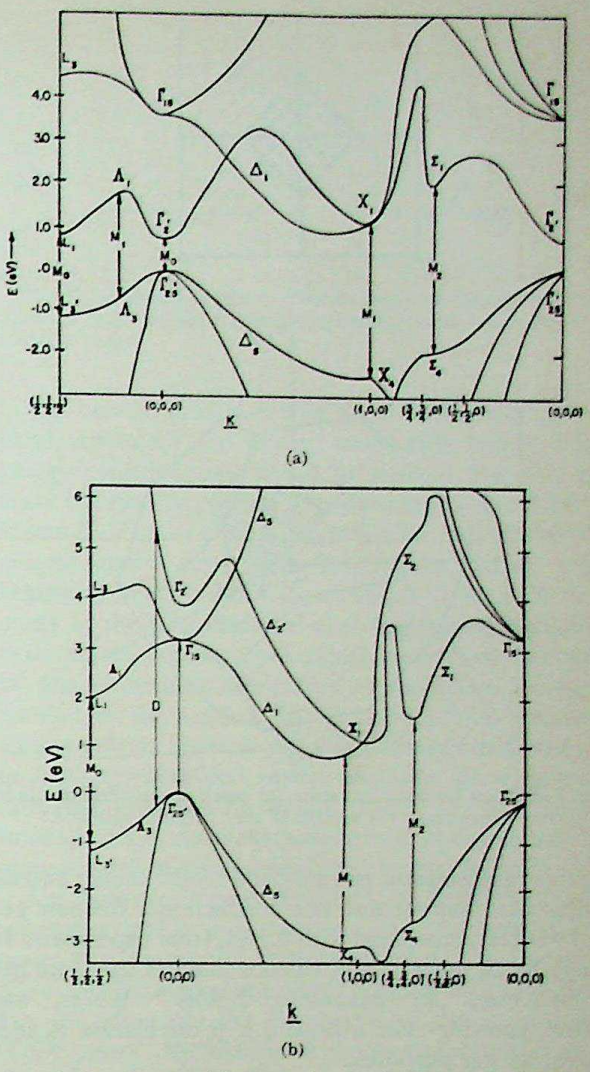


FIG. 6. Pseudopotential energy bands along Δ, Λ, and Σ symmetry directions. Some of the principal transitions have been marked.

wave functions the momentum matrix element was found to be $0.73\hbar^2(2\pi/a)^2$. The energy denominator is 1.8 eV. Putting these values into (18) one finds $m_t^*=0.12m_e$ compared with the experimental value of $0.082m_e$. Considering the approximations, used to get m_t^* the result is reasonably close to experiment. This value was checked by directly computing $E(\mathbf{k})$ in a transverse direction and using

$$m_e/(m_t^*)=[2m_e(E(\Delta\mathbf{k}+\mathbf{k}_L)-E(\mathbf{k}_L))]/\hbar^2(\Delta\mathbf{k})^2.$$

This also gave us $0.12m_e$, which justifies the use of a two-band model in the $\mathbf{k}\cdot\mathbf{p}$ calculation.

For Si where the minima correspond to Δ_1 states, we have from Dresselhaus[33]

$$\left(\frac{m_e}{m_t^*}\right)=1+\sum_1\frac{|\langle\Delta_1|P_y|\Delta_5{}^1\rangle|^2}{(E_{\Delta_1}-E_{\Delta_5{}^1})}. \qquad (19)$$

[33] G. F. Dresselhaus, Ph.D. thesis, University of California (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

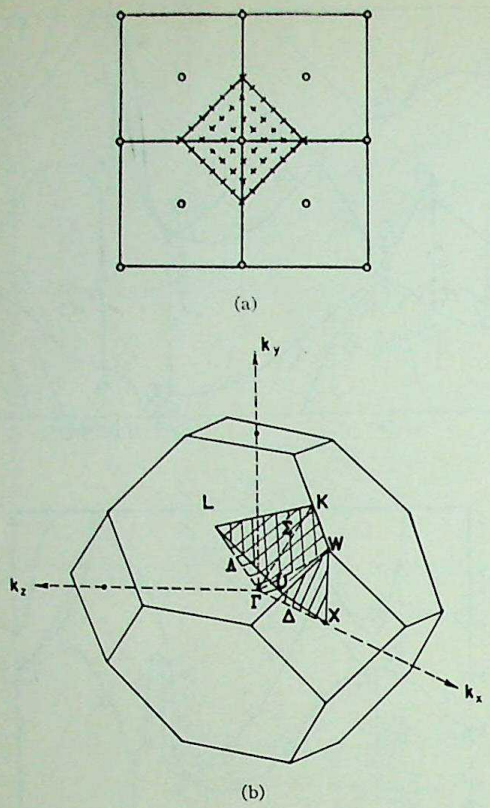


FIG. 7. Scheme for sampling zone (a) mesh of sampling points in two dimensions, (b) section of zone actually sampled.

In this case we must use a 3-band model taking two Δ_5 states (one valence and one conduction). We now get $m_t^*=0.17m_e$ compared with $0.19m_e$ from experiment. If only the Δ_5 state in the valence band is used, we get $m_t^*=0.14m_e$. By examining $E(\mathbf{k})m_t^*=0.17m_e$ was found, showing that a 3-band $\mathbf{k}\cdot\mathbf{p}$ calculation is adequate for our purposes.

VII. SAMPLING PROCEDURES

With a view toward finding $\epsilon_2(\omega)$ we discuss the calculation of the joint density-of-states functions $J_{ns}(\omega)$ defined by Eq. (2). Since in practice it is possible to solve for the eigenvalues of H_p at only a finite number of points in the reduced zone, we must replace the integral in (2) by a finite sum. This was done by defining the quantity

$$K_{ns}(\omega_i)=\sum_{\mathbf{k}_l}\delta^{\Delta\omega}(\omega_{ns}(\mathbf{k}_l)-\omega_i). \tag{20}$$

Here $\mathbf{k}_l$ is a set of uniformly spaced sampling points lying within the first Brillouin zone. The delta function is defined by

$$\delta^{\Delta\omega}(\omega_{ns}-\omega_0)=1 \quad \text{if} \quad |\omega_{ns}-\omega_0|\leqslant(\Delta\omega/2)=0 \text{ otherwise.} \tag{21}$$

The function $K_{ns}(\omega_i)$ is then defined for a set of values ω_i such that $\omega_{i+1}=\omega_i+\Delta\omega$. The meaning of $K_{ns}(\omega_i)$ is clear. It measures the number of pairs of states in bands n and s with direct energy gap $\hbar\omega_{ns}=\hbar\omega_i\pm\hbar(\Delta\omega/2)$ as found in the sample. The relationship between $K_{n,s}(\omega_i)$ and $J_{ns}(\omega)$ is

$$K_{ns}(\omega_i)=\frac{\Delta\omega}{\Delta^3k}\frac{(2\pi)^3}{2}J_{ns}(\omega=\omega_i), \tag{22}$$

where Δ^3k is the elementary volume surrounding the sampling points.

In making a choice for $\Delta\omega$ we were guided by the consideration that most of the experimental structure in $\epsilon_2(\omega)$ has a half-peak width $\sim\frac{1}{2}$ eV. We consequently took $\Delta\omega=0.1$ eV$/\hbar$ as a reasonable value.

The reciprocal lattice vectors of the diamond structure form a bcc lattice. The most natural choice for the sampling vectors $\mathbf{k}_l$ is a mesh of points forming a bcc array. This was reduced in size from the host lattice by a scaling factor of 36. That is, the $\mathbf{k}_l$ are determined by $\mathbf{k}_l=\mathbf{K}_l/36$ where $\mathbf{K}_l$ are reciprocal lattice vectors. This idea is demonstrated for a two-dimensional body centered structure in Fig. 7(a). With this reduction in scale the calculation gives results for $36^3=46\,656$ $\sim 50\,000$ points in the reduced zone.

Advantage was taken of the 48-fold symmetry of the Brillouin zone of the diamond structure. It is then necessary to examine only that 1/48th part of the zone indicated in Fig. 7(b). This region is defined by the conditions

$$0\leqslant(\mathbf{k}_l)_z\leqslant(\mathbf{k}_l)_y\leqslant(\mathbf{k}_l)_x, \tag{23}$$

and $\mathbf{k}_l$ contained within the first Brillouin zone. The contribution to $K_{ns}(\omega_i)$ from a point $\mathbf{k}_l$, in the interior of the sampling region is then taken to be 48. The number 48 is referred to as the weighting factor W of the point $\mathbf{k}_l$.

For points not in the interior of the sampling region W is easily found. We simply find the number of regions between which the point $\mathbf{k}_l$ is shared. If we call this number C, then $W=48/C$. For points not on the Brillouin zone surface C is found as follows:

$$\begin{aligned}
C&=\left\{\begin{matrix}1\\2\end{matrix}\right\}, \quad \text{if} \quad (\mathbf{k}_l)_x\neq(\mathbf{k}_l)_y\neq(\mathbf{k}_l)_z\begin{matrix}\neq0\\=0\end{matrix}\Big\},\\
C&=\left\{\begin{matrix}2\\8\end{matrix}\right\}, \quad \text{if} \quad (\mathbf{k}_l)_x\neq(\mathbf{k}_l)_y=(\mathbf{k}_l)_z\begin{matrix}\neq0\\=0\end{matrix}\Big\},\\
C&=\left\{\begin{matrix}2\\4\end{matrix}\right\}, \quad \text{if} \quad (\mathbf{k}_l)_x=(\mathbf{k}_l)_y\neq(\mathbf{k}_l)_z\begin{matrix}\neq0\\=0\end{matrix}\Big\},\\
C&=\left\{\begin{matrix}6\\48\end{matrix}\right\}, \quad \text{if} \quad (\mathbf{k}_l)_x=(\mathbf{k}_l)_y=(\mathbf{k}_l)_z\begin{matrix}\neq0\\=0\end{matrix}\Big\}.
\end{aligned} \tag{24}$$

If $\mathbf{k}_l$ lies on the first Brillouin zone surface then it is shared with adjoining zones, and the above value of C is multiplied by 2, 3, or 4 according to whether $\mathbf{k}_l$ is on a face, edge, or corner. For example, for the point Γ, $C=48$ and $W=1$. With this technique H_p had only to be diagonalized at ~ 1200 points.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

When the data first came from the computer, it was observed that the scatter in the 0.1-eV histograms, $K_{ns}(\omega_i)$, was excessive. We found, however, that by smoothing $K_{ns}(\omega_i)$ according to

$$\bar{K}_{ns}(\omega_i)=\tfrac{1}{3}[K_{ns}(\omega_{i+1})+K_{ns}(\omega_i)+K_{ns}(\omega_{i-1})]; \quad (25)$$

the scatter was materially reduced. This method is discussed by Hartree.[34] When the smoothing procedure is combined with the results of critical point analysis as in the next section, excellent results are achieved. In the Appendix we examine the effects of smoothing as well as the sources of the statistical noise.

VIII. DISCUSSION OF LINE SHAPES

In discussing the results we shall begin by examining the histogram for bands 4 and 5 (the bands are ordered in the conventional way with $n=4$ the highest valence band, and $s=5$ the lowest conduction band). From energy considerations it can be guessed that this will make the most important contribution to $\epsilon_2(\omega)$. Figure 8 shows $\bar{K}_{4,5}(\omega)$ for both Ge and Si. One notices that the graphs show a group of periodic peaks (dashed lines) beginning at 4.5 eV in Ge and at 5.6 eV in Si. A close study of the data disclosed that these arise in an artificial way as boundary sampling errors. Near the point W one finds that the planes of sampling points lie parallel to the surfaces on which $E_{4,5}(\mathbf{k})$ is constant. This condition is drawn for a two-dimensional example in Fig. 9, where one easily sees how these artificial sampling peaks can be formed. Since they were judged to be spurious, the peaks have been eliminated from the joint density of states (solid lines in Fig. 8).

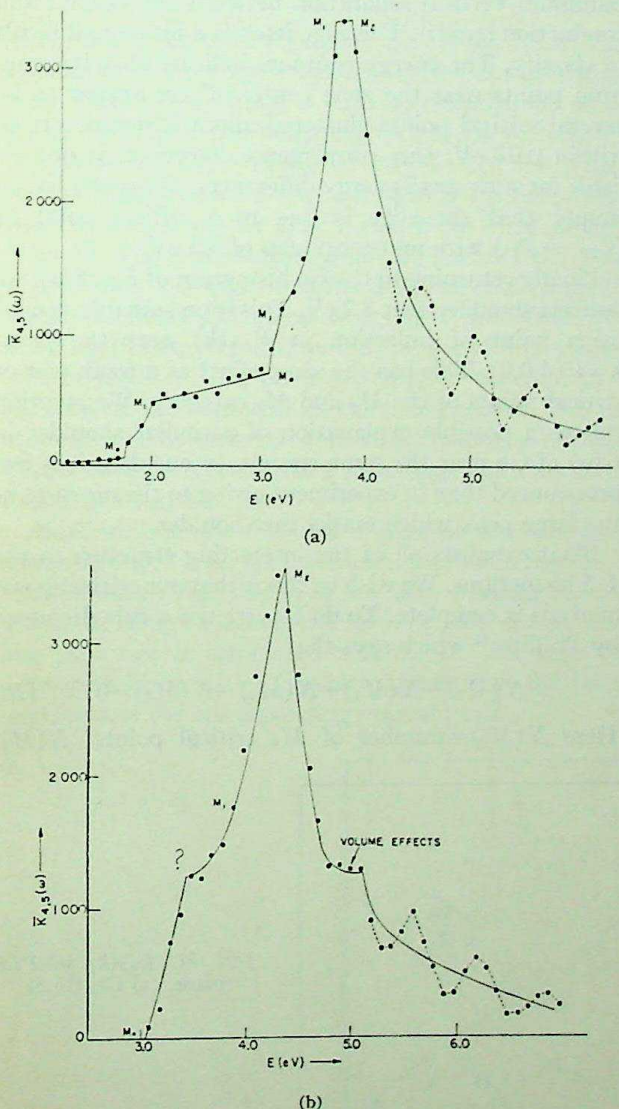


FIG. 8. 4–5 histograms. Volume effects refer to structure which could not be traced to critical points. (a) Ge, (b) Si.

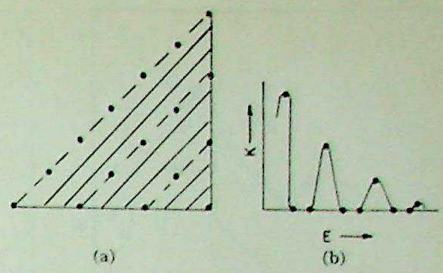


FIG. 9. (a) Sampling planes (dashed lines) are parallel to energy contours (solid lines); (b) resulting density of states showing the spurious peaks.

The most striking result is the large central peak. We remember that in our starting model the large optical peak was associated with transitions of the type $X_4 \rightarrow X_1$. To understand how it arises in $\bar{K}_{4,5}$ we inspect Fig. 10 where $E_{4,5}(\mathbf{k})$ is plotted in the basal plane (ΓKX plane). By looking at the energy contours near the point X in Fig. 10(b), one discovers that there exists a critical point of the M_1 class with a direct energy gap of 3.6 eV. Furthermore, on the Σ axis near the point $\mathbf{k}=(0.6,0.6,0.0)$ there occurs another critical point. This belongs to the M_2 class, and has an energy gap of 3.8 eV. The large peak in $\bar{K}_{4,5}$ of Ge, results then from these two nearly degenerate critical points corresponding to $X_4 \rightarrow X_1$ and $\Sigma_4 \rightarrow \Sigma_1$ transitions. This idea has been demonstrated previously in Fig. 1(e). The occurrence of a critical point on Σ is a new result.

The analysis of Fig. 10(b) for Si leads to the same general conclusions. At the point X there again occurs an M_1 critical point with energy of 4.0 eV. On the Σ axis near $\mathbf{k}=(0.4,0.4,0.0)$ we find an M_2 critical point this time with an energy gap of 4.4 eV.

In Fig. 8 the reader may notice that sharp corners have been drawn into the data. The analysis discussed earlier indicates that this is the correct behavior near a critical point.

Returning to Fig. 8(a), we see a sharp corner of the M_1 type in $\bar{K}_{4,5}$ of Ge at 2.0 eV. It is clear that this is identifiable with the spin-orbit split optical edge at 2.2 eV. Figure 11(a) shows the contours of $E_{4,5}(\mathbf{k})$ in the

[34] D. R. Hartree, *Numerical Analysis* (Oxford University Press, London, 1952), pp. 249–251.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

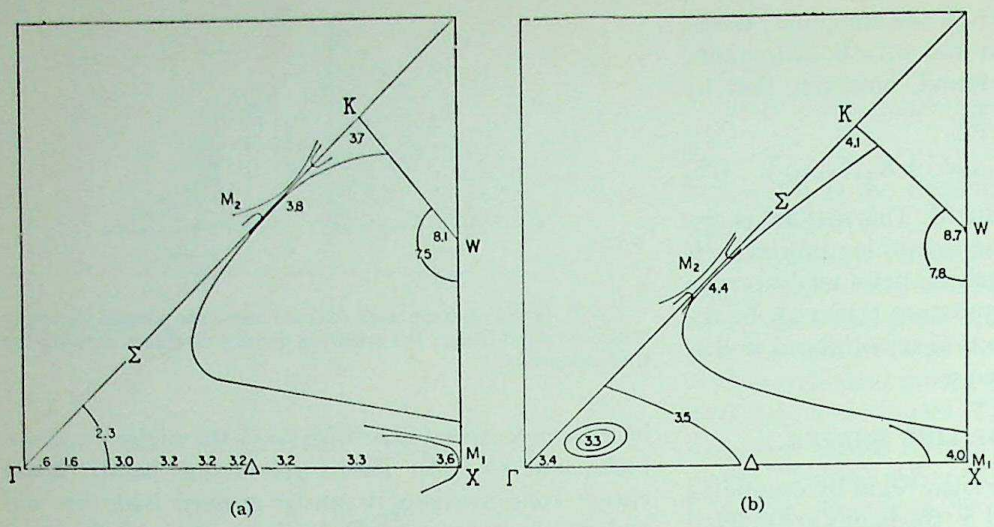


FIG. 10. $E_{4,5}(\mathbf{k})$ contours in ΓKX plane. The energies indicated on graph are all in eV. A few of the principal contours are drawn showing behavior near critical points. (a) Ge, (b) Si.

plane defined by ΓKL. At L there is an M_0 critical point with a gap of 1.8 eV. Looking along Λ, we detect another critical point near $\mathbf{k}=(0.17,0.17,0.17)$. This one is classified as an M_1 with an energy of 2.0 eV. The edge results from $\Lambda_3 \rightarrow \Lambda_1$ transitions, while the threshold is produced by $L_{3'} \rightarrow L_1$ transitions. This interpretation is again a new one.

Both $\Lambda_3 \rightarrow \Lambda_1$ and $L_{3'} \rightarrow L_1$ are spin-orbit split by approximately the same amount. Although the predicted breadth of the 2-eV edge given by the separation of the M_0 and M_1 singularities (0.2 eV) is in good agreement with experiment, more decisive confirmation of the predicted M_0 threshold can be obtained in materials where the spin-orbit splitting and the M_1 peak are somewhat larger. This is the case for GaAs[35] and still more clearly ZnTe.[36] The M_0 thresholds are well resolved approximately 0.2 eV below the M_1 edges. It is striking that such fine details appear to carry over not only to slightly ionic crystals such as GaAs but also to strongly ionic ones such as ZnTe.

In the 4–5 histogram of Si there is an edge at 3.5 eV. This is identified with the peak at 3.5 eV in $\epsilon_2(\omega)$. Figure 11(b) shows that the threshold for this edge is once more an M_0 critical point at L. The energy of the $L_{3'} \rightarrow L_1$ gap is here 3.1 eV (in Si this appears to be the minimum vertical separation between the valence and conduction bands). The edge itself is a bit more difficult to classify. The energy contours indicate that it comes from points near the zone center. There appear to be several critical points clustered about Γ degenerate to within 0.04 eV. Our convergence, however, is not reliable for such small energy differences. We prefer to say simply that the edge is due to a critical point at $(\Gamma_{25'} \rightarrow \Gamma_{15})$ with an energy gap of 3.5 eV.

Finally returning to the Ge histogram of Fig. 8(a), we notice a shoulder near 3.2 eV. This is presumably caused by a point of inflection in $E_{4,5}(\mathbf{k})$ near the point $\mathbf{k}=(0.4,0,0)$. This has the same effect as a weak pair of critical points of the M_0 and M_1 category. We mention this as a possible explanation of a similar shoulder in $\epsilon_2(\omega)$ of Ge near the same energy. In our data it is less pronounced than in experiment owing to the nearness of the large peak which masks the shoulder.

That exhausts all of the interesting structure in the 4–5 histograms. We wish to check that our critical-point analysis is complete. To do this we use a rule discussed by Phillips,[25] which says that

$$N(M_0)-N(M_1)+N(M_2)-N(M_3)=0. \qquad (26)$$

Here $N(M_0)$=number of M_0 critical points; $N(M_1)$

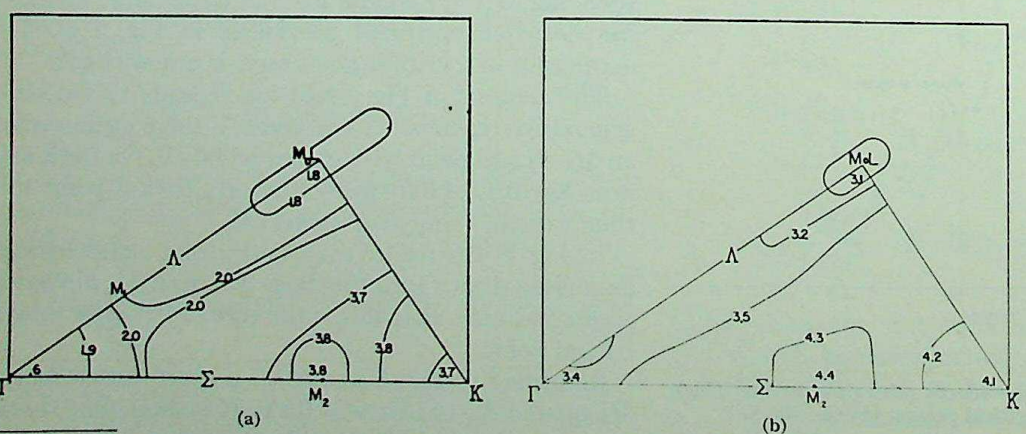


FIG. 11. $E_{4,5}(\mathbf{k})$ in ΓLK plane. (a) Ge, (b) Si.

[35] D. L. Greenaway, Phys. Rev. Letters 9, 97 (1962).
[36] M. Cardona and D. L. Greenaway, Phys. Rev. 131, 98 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

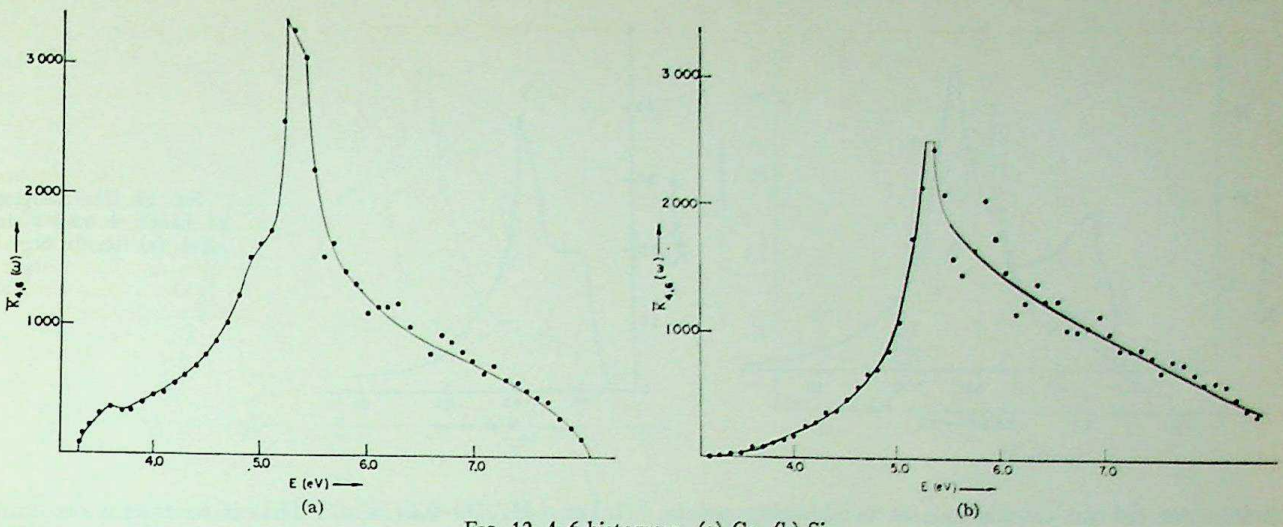


FIG. 12. 4–6 histogram. (a) Ge, (b) Si.

=number of M_1 critical points; etc. In Ge, $N(M_0)$ has contributions from three sources: Γ (the absolute minimum $\Gamma_{25'} \rightarrow \Gamma_{2'}$), L, and Δ (the M_0 contribution of the kink). Putting in the weighting factors (W) for each of the points we have $N(M_0)=1+4+6=11$. For $N(M_1)$ the contributions are from Λ, X, and Δ (the M_1 saddle point arising from the kink). This gives $N(M_1)=8+3+6=17$. The contributions to $N(M_2)$ are all from Σ and $N(M_2)=12$. The contributions to M_3 all come from the maximum in $E_{4,5}(\mathbf{k})$ which occurs at W and $N(M_3)=6$. Forming the sum implied by (26) we get $11-17+12-6=0$ as required. We have not applied the rule to Si because of the complexities discussed earlier. We can nevertheless, be confident that all the important critical points have been found since the energy contours in Ge and Si are very similar.

Next we examine the contribution of the 4–6 histogram. In Fig. 12, $\bar{K}_{4,6}(\omega)$ is drawn. The main feature in the Ge data [Fig. 12(a)] is the large peak at 5.3 eV. From the energy contours of Fig. 13(a), one can see that the structure is associated with L. In this case the transition being $L_{3'} \rightarrow L_3$ which agrees with the interpretation of the corresponding optical peak.

The Si data are understood similarly. The $L_{3'} \rightarrow L_3$ transition is responsible for the peak at 5.3 eV. The somewhat large peak at 5.8 eV is apparently a boundary sampling error of the type discussed in connection with the 4–5 histogram. After matrix elements are accounted for, the remaining structure in $\bar{K}_{4,6}$ is too weak to make an observable contribution to the optical properties. For the same reason the remaining histograms contribute nothing to $\epsilon_2(\omega)$.

We are now in a position to use these results to construct $\epsilon_2(\omega)$. First we shall give a very brief discussion of how the matrix elements were determined.

In order to develop a reasonable way of handling $|M_{n,s}(\mathbf{k})|^2$ we computed matrix elements for $n=4$, $s=5$, and for $n=4$, $s=6$, at a number of points in the reduced zone. These were gotten by using the pseudopotential wave functions $\phi_{n,\mathbf{k}}$. Such a procedure corresponds to using the smooth part of the OPW wave functions which according to Ref. 37 gives accuracy of better than 20%.[38] The results are shown in Table III. The Si wave functions were used; however, we would get almost identical results with the Ge wave functions. In obtaining the matrix elements for the symmetry

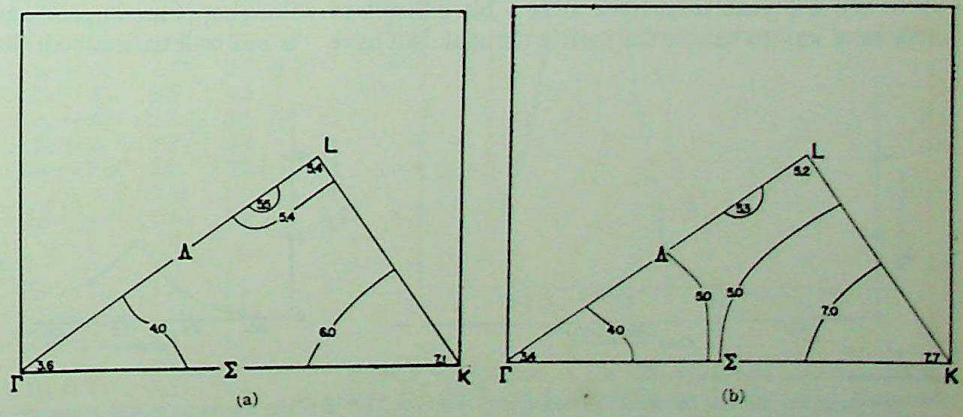


FIG. 13. $E_{4,6}(\mathbf{k})$ in ΓLK plane. (a) Ge, (b) Si.

[37] L. Kleinman and J. C. Phillips, Phys. Rev. 118, 1153 (1960).
[38] The effective masses are determined by momentum matrix elements. Since our computed values for the effective masses agree well with experiment, we naturally expect the computed matrix elements also to be correct.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

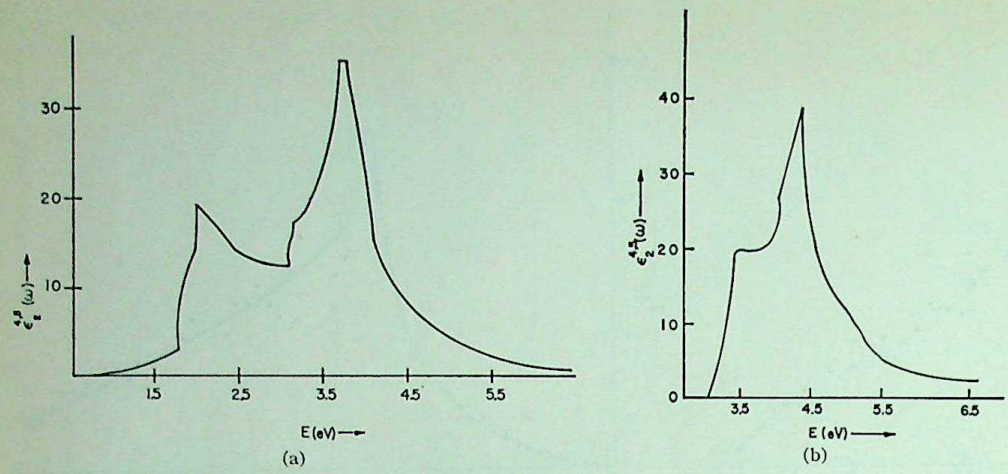


FIG. 14. Contribution of bands 4 and 5 to $\epsilon_2(\omega)$. (a) Ge, (b) Si.

points, we did the calculations at neighboring points. This was done so that degeneracies would be lifted, and the resulting matrix elements would be those of the bands being studied. If this were not done then, for example, the wave functions of bands 5 and 6 would mix at X.[39]

Looking at the values for $|M_{4,5}(\mathbf{k})|^2$ in Table III we see that it is roughly constant throughout the zone

TABLE III. Matrix elements in units of $(2\pi/a)^2$.

k	$\lvert M_{4,5}(\mathbf{k})\rvert^2$	$\lvert M_{4,6}(\mathbf{k})\rvert^2$
(0.04,0.04,0.02)	0.90	0.01
(1.00,0.04,0.04)	1.51	0.00
(0.72,0.39,0.33)	1.33	0.09
(0.6,0.6,0.0)	1.41	0.00
(0.58,0.25,0.14)	0.94	0.22
(0.50,0.48,0.52)	0.85	0.63

varying by less than a factor of 2. We could then replace the matrix element in Eq. (1) by a $\mathbf{k}$-independent quantity $|M_{4,5}|^2$. To determine this value we took the weighted average of the values in Table III. Then $|M_{4,5}|^2=1.2(2\pi/a)^2$. As Table III points out, $|M_{4,6}(\mathbf{k})|^2$ is small except near the point L, precisely where the 4–6 peak originates. It is a bit difficult to settle on a way to handle the matrix element. We have taken $|M_{4,6}|^2=0.6(2\pi/a)^2$. That is we take a constant value equal to $|M_{4,6}(\mathbf{k}=(\frac{1}{2},\frac{1}{2},\frac{1}{2}))|^2$. In Figs. 14 and 15 we show the contributions to $\epsilon_2(\omega)$ of the 4–5 and 4–6 histograms. In Fig. 16 we take $\epsilon_2(\omega)=\epsilon_2^{4,5}(\omega)+\epsilon_2^{4,6}(\omega)$, that is, as the sum of the contributions from the 4–5 and 4–6 histograms, and compare with experiment. In previous work (see Refs. 40 and 41) $\epsilon_2(\omega)$ was also taken as $=\epsilon_2^{4,5}(\omega)+\epsilon_2^{4,6}(\omega)$. Figure 6 shows our energy bands with the principal transitions marked.

To check our hypothesis regarding matrix elements it was decided to do Si with matrix elements included. $|M_{n,s}(\mathbf{k})|^2$ was computed along with the energy eigenvalues at each point of the zone, and put in Eq. (1). This added about 3 h of machine time to a total of about 15 h. The result for $\epsilon_2(\omega)$ shown in Fig. 16(b) includes all of the separate histograms.

As Fig. 16 indicates, our line shapes appear satisfactory. The two exceptions are the $L_{3'}\rightarrow L_3$ peaks in Ge and Si, as well as the $\Gamma_{25'}\rightarrow\Gamma_{15}$ peak in Si. The main difficulty in the first case centers around the strength. We have seen that our treatment probably overestimates the matrix element so that it is not surprising that the computed strength is too large. In Fig. 17 we examine the experimental line shape derived by taking the shaded portion of $\epsilon_2(\omega)$ as shown by the inset. With suitable magnification the theoretical and experimental line shapes are in good agreement. The $\Gamma_{25'}\rightarrow\Gamma_{15}$ peak is not well understood. Since it is so narrow, there is a

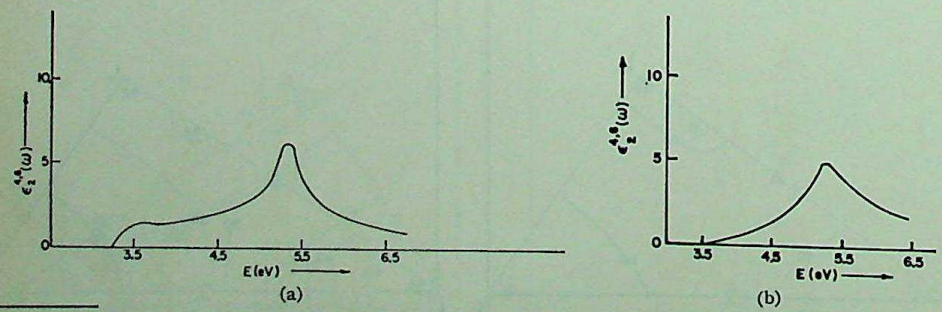


FIG. 15. Contribution of bands 4 and 6 to $\epsilon_2(\omega)$. (a) Ge, (b) Si.

[39] The program did not construct symmetrized combinations of plane waves as it was a general one designed to handle all points of the zone on an equal footing.

[40] D. Brust, J. C. Phillips, and F. Bassani, Phys. Rev. Letters 9, 94 (1962).

[41] D. Brust, M. L. Cohen, and J. C. Phillips, Phys. Rev. Letters 9, 389 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

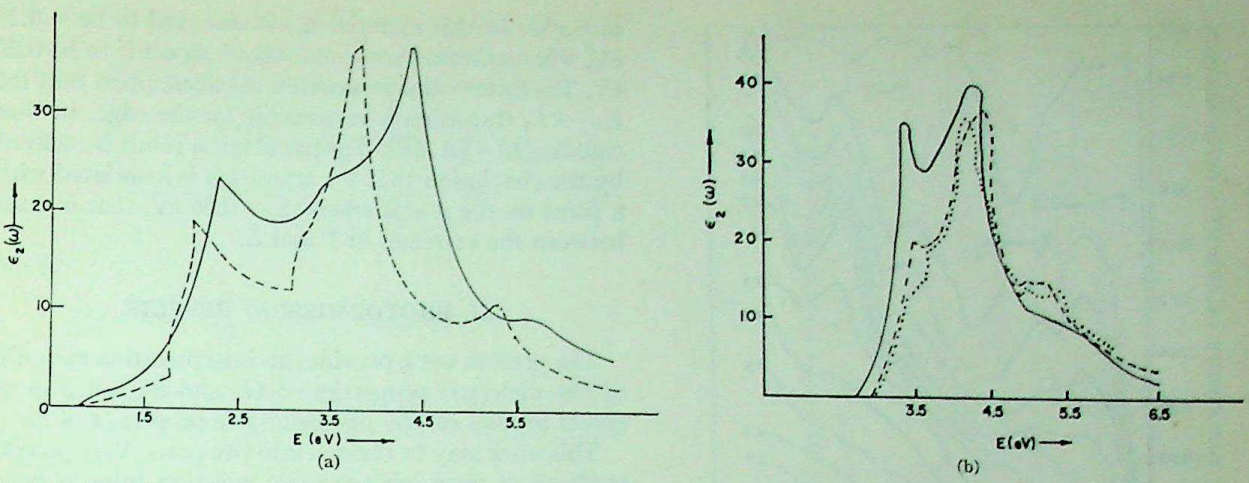


FIG. 16. Experimental result for $\epsilon_2(\omega)$ (solid line) compared with our computed value (a) Ge $\epsilon_2(\omega)=\epsilon_2^{4,5}(\omega)+\epsilon_2^{4,6}(\omega)$ (dashed line, with edges emphasized to account for critical points), (b) Si $\epsilon_2(\omega)=\epsilon_2^{4,5}(\omega)+\epsilon_2^{4,6}(\omega)$ (dashed line), $\epsilon_2(\omega)$ taken as sum of contribution from all bands with dipole matrix elements included (dotted line).

chance that it is a density-of-states effect which we did not detect. The reader should see a further discussion concerning this point by Phillips.[42]

The results of this section are summarized in Table IV.

IX. STATIC DIELECTRIC CONSTANT

From a knowledge of $\epsilon_2(\omega)$ it is possible to calculate the real part of the dielectric constant $\epsilon_1(\omega)$ in the limit $\omega \rightarrow 0$. From the Kramers-Kronig relation the following expression can be obtained[43]

$$\epsilon_1(0)=1+\frac{2}{\pi}\int_0^\infty \frac{\epsilon_2(\omega)}{\omega}d\omega.$$

TABLE IV. Important critical points below 6 eV. In some cases they have been given to 0.01 eV with a reliability ~0.03 eV.

Type	Location	Bands	Symbol	Theoretical energy	Experimental energy
		Germanium			
M_0	(0,0,0)	$4\rightarrow 5$	$\Gamma_{25'}\rightarrow\Gamma_{2'}$	0.6	0.8
M_0	(0.5,0.5,0.5)	$4\rightarrow 5$	$L_{3'}\rightarrow L_1$	1.78	2.1
M_1	(0.17,0.17,0.17)	$4\rightarrow 5$	$\Lambda_3\rightarrow\Lambda_1$	2.01	2.3
M_1	(0.3,0.0,0.0)	$4\rightarrow 5$	$\Delta_5\rightarrow\Delta_1$	3.17	3.0 (?)
M_1	(1.0,0.0,0.0)	$4\rightarrow 5$	$X_4\rightarrow X_1$	3.6	4.3
M_2	(0.5,0.0,0.0)	$4\rightarrow 5$	$\Delta_5\rightarrow\Delta_1$	3.21	3.4 (?)
M_2	(0.61,0.61,0.0)	$4\rightarrow 5$	$\Sigma_4\rightarrow\Sigma_1$	3.8	4.4
M_0	(0.0,0.0,0.0)	$4\rightarrow 6$	$\Gamma_{25'}\rightarrow\Gamma_{15}$	3.6	3.2 (?)
M_1	(0.56,0.56,0.39)	$4\rightarrow 6$	...	5.33	5.6
M_2	(0.5,0.5,0.5)	$4\rightarrow 6$	$L_{3'}\rightarrow L_3$	5.44	5.7
		Silicon			
...	(0.0,0.0,0.0)	$4\rightarrow 5$	$\Gamma_{25'}\rightarrow\Gamma_{15}$	3.5	3.5
M_0	(0.5,0.5,0.5)	$4\rightarrow 5$	$L_{3'}\rightarrow L_1$	3.15	3.0
M_1	(1.0,0.0,0.0)	$4\rightarrow 5$	$X_4\rightarrow X_1$	4.0	4.3
M_2	(0.4,0.4,0.0)	$4\rightarrow 5$	$\Sigma_4\rightarrow\Sigma_1$	4.4	4.4
M_0	(0.5,0.5,0.5)	$4\rightarrow 6$	$L_{3'}\rightarrow L_3$	5.2	5.4

[42] J. C. Phillips, Phys. Rev. Letters **10**, 329 (1963).
[43] T. S. Moss, Proc. Phys. Soc. (London) **B66**, 141 (1953).

Using our results for $\epsilon_2(\omega)$, we can now derive $\epsilon_1(0)$. It is realized, however, that the value calculated in this way is directly proportional to the choice of matrix elements. In the case of the 4–5 bands we took this as $1.2(2\pi/a)^2$ with an assigned error ~25%. For the other pairs of bands where the matrix element is considerably smaller we do not have a sufficiently reliable value for this purpose. Therefore, only the contribution from $\epsilon_2^{4,5}(\omega)$

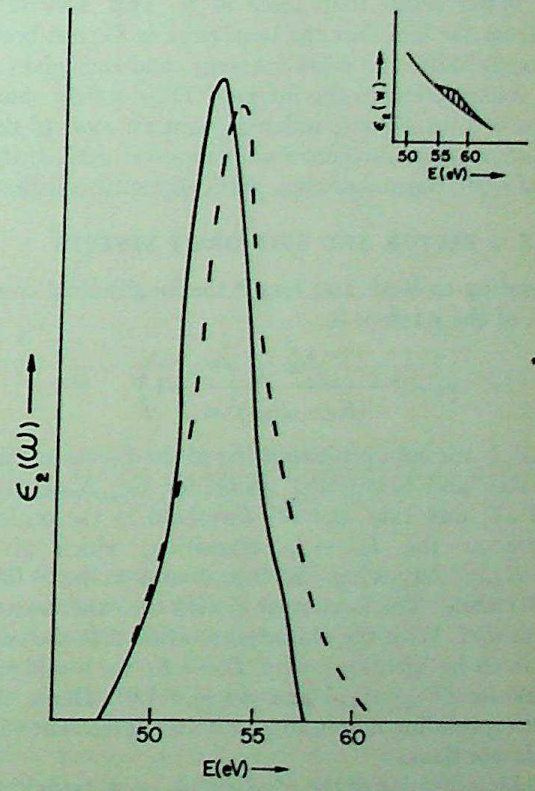


FIG. 17. Comparison of the theoretical and experimental line shapes for the $L_{3'}\rightarrow L_3$ transition in Ge. The solid line is the shaded portion of the experimental curve indicated by the inset. The shaded region has been moved 0.5 eV to the left and magnified seven times. The dashed line is a similar region under the theoretical curve.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

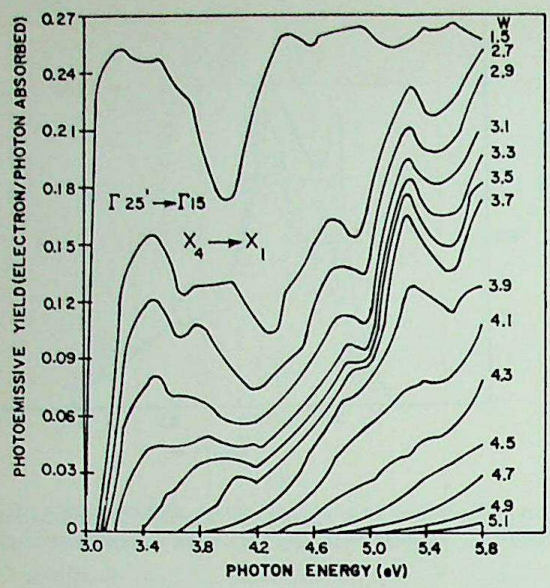


FIG. 18. Photoemissive yield: electrons/absorbed photons, as a function of photon energy, with vacuum W, eV, above the top of the valence band, as computed in present work.

was evaluated. For Ge this gave $\epsilon_1^{4,5}(0)=12.4$ and for Si $\epsilon_1^{4,5}(0)=7.6$, which are to be compared with the experimental values of 16.0 and 12.0, respectively. As expected $\epsilon_1^{4,5}(0)<\epsilon_1(0)$ due to the neglect of the higher bands. The 4–5 bands of Ge appear from these values to give a better result than those of Si. This, however, arises from the fact that the large peak in Ge has been erroneously shifted to a lower energy, and thus gives a bigger contribution to the integral. These values show that the 4–5 transitions, which account for most of the structure in $\epsilon_2(\omega)$, also contribute more than 50% of the value of $\epsilon_1(0)$, the macroscopic static dielectric constant.

X. g FACTOR AND SPIN-ORBIT EFFECTS

According to Roth and Lax,[15] the longitudinal component of the g tensor is

$$g_{11}=2-\frac{\Delta_{s0}}{\langle E_{L_1}-E_{L_3'}\rangle}\left(\frac{m_0}{m_t}-1\right).$$

Here Δ_{s0} is the spin-orbit splitting of the $L_{3'}$ states. We use Lukěs and Schmidt's[44] value for Ge, $\Delta_{s0}=0.195\pm0.02$ eV, and take the M_0 threshold in the optical response as the $L_{3'}\rightarrow L_1$ transition, which gives $\langle E_{L_1}-E_{L_3'}\rangle=2.1$ eV. g_{11} is then found to be $=0.96$ ($m_t=0.082m_0$). This is compared with the experimental value of 0.87. With the old interpretation that the edge itself is to be identified with $L_{3'}\rightarrow L_1$ we would use 2.35 eV for $\langle E_{L_1}-E_{L_3'}\rangle$ and get $g_{11}=1.07$. Hence the new interpretation is seen to give better agreement with the g-factor data.

The identification of the edge with $\Lambda_3\rightarrow\Lambda_1$ transitions allows us to resolve a problem in the spin-orbit splitting of GaAs. In this material Δ_{s0} is observed to be $=0.26$ eV, whereas from theory one would expect it to be 0.20 eV. The latter value is based on the assumption that the $L_{3'}\rightarrow L_1$ transition is responsible for the edge. In that case $\Delta_{s0}(L)=\frac{2}{3}\Delta_{s0}(\Gamma)$. The paradoxical result is resolved by the conclusion that the transition is associated with a point on the Λ axis where $\Delta_{s0}=0.26$ eV, that is, falls between the extremes at Γ and L.

XI. PHOTOEMISSION RESULTS

The present work provides an interpretation not only of the dielectric properties of Ge and Si, but also of recent studies on the photoemissive properties of Si.

This work may be divided into two parts. Very precise studies[45] of weak photoelectric emission from Si near threshold on atomically clean surfaces reveal direct and indirect thresholds similar to the infrared absorption edges discussed by Hall, Bardeen, and Blatt.[46] The indirect threshold is at 5.15 eV, the direct threshold at 5.45 eV. The crystal surface is a (111) plane, and the thresholds, according to Kane,[47] are associated with the symmetry line Λ, where V is directed normal to the (111) plane. The direct transition responsible for the direct threshold therefore starts *from* Λ_3 at an energy 0.3 eV below $\Gamma_{25'}$. The transition marked D in Fig. 6(b) starts from $\mathbf{k}=(0.12,0.12,0.12)$ and requires $E=5.5$ eV. The agreement with Kane's theory and Allen and Gobelli's experiment is satisfying. It should be emphasized that this agreement is a by-product of our calculations for the ultraviolet absorption.

The second part of the photoelectric data concerns the broad yield curves obtained when the effective work function is reduced by covering the surface with up to one monolayer of Cs. In doing the analysis we assume that electrons which are excited into the conduction bands have an escape probability $P_s^W(\mathbf{k})$. Then $P_s^W(\mathbf{k})=0$ if $E_s(\mathbf{k})<W$ and $P_s^W=P$ if $E_s(\mathbf{k})>W$. Here $E_s(\mathbf{k})$ is the energy of the electron state with respect to the top of the valence band and W is the vacuum energy also measured relative to $\Gamma_{25'}$. With these assumptions the photoemissive yield function (number of electrons emitted/number of photons absorbed) is given by

$$Y_W(\omega)=\frac{\sum_{n,s}\int_{\text{B.Z.}} d^3k P_s^W(\mathbf{k})\delta(\omega_{n,s}(\mathbf{k})-\omega)\,|M_{n,s}(\mathbf{k})|^2}{\sum_{n,s}\int_{\text{B.Z.}} d^3k\,\delta(\omega_{n,s}(\mathbf{k})-\omega)\,|M_{n,s}(\mathbf{k})|^2},\tag{27}$$

in which it is assumed that the electron escapes with a negligible loss of energy.

[44] F. Lukěs and E. Schmidt, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 389.

[45] G. W. Gobelli and F. G. Allen, Phys. Rev. 127, 141 (1962).

[46] L. H. Hall, J. Bardeen, and F. J. Blatt, Phys. Rev. 95, 559 (1954).

[47] E. O. Kane, Phys. Rev. 127, 131 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In performing the calculations we arbitrarily set $P=0.26$. Then, again, taking constant matrix elements we have after making a finite sample of the zone

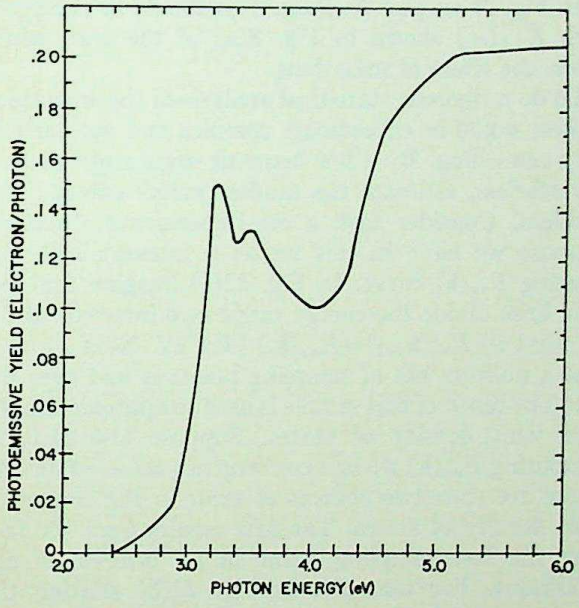
$$Y_W(\omega_i)=\frac{\sum_{n,s} K_{n,s}{}^W(\omega_i)\,|M_{n,s}|^2}{\sum_{n,s} K_{n,s}(\omega_i)\,|M_{n,s}|^2}, \qquad (28)$$

where

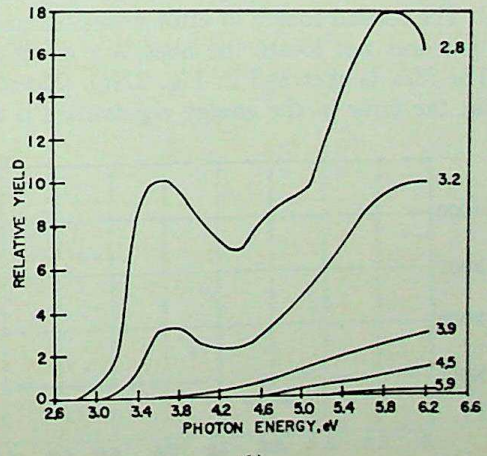
$$K_{n,s}{}^W(\omega_i)=\sum_{\mathbf{k}_l} P_s{}^W(\mathbf{k}_l)\hat{\delta}^{\Delta\omega}(\omega_{n,s}(\mathbf{k}_l)-\omega_i). \qquad (29)$$

Again we replace the quantities $K_{n,s}(\omega_i)$ and $K_{n,s}{}^W(\omega_i)$ by the three point smoothed quantities $\bar{K}_{n,s}(\omega_i)$ and $\bar{K}_{n,s}{}^W(\omega_i)$.

In Fig. 18 we have plotted the results for $Y_W(\omega)$ with ω varied between 1.5 and 5.1 eV. This is to be compared

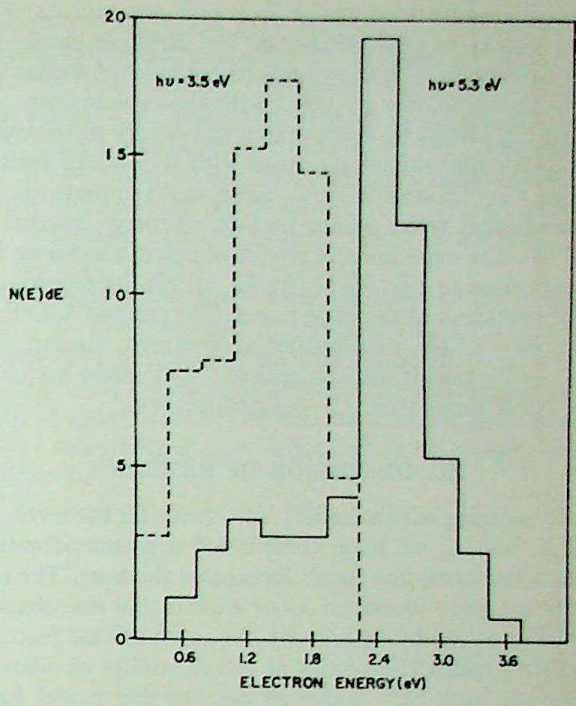


(a)

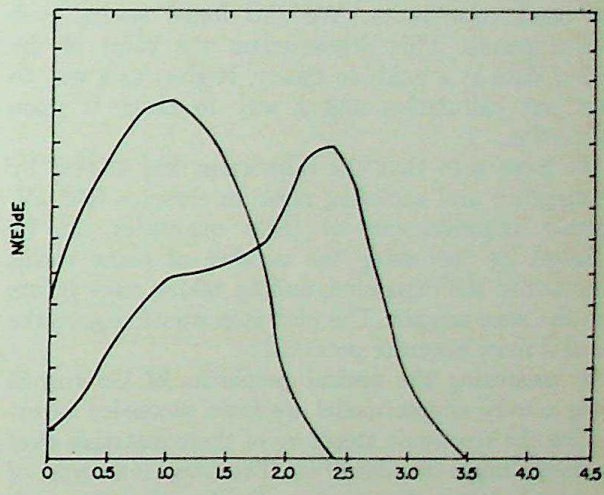


(b)

FIG. 19. Experimental results for the photoemissive yield (a) Spicer; (b) Gobeli and Allen.



(a)



(b)

FIG. 20. (a) The energy distribution of electrons emitted from Si for $\hbar\omega=3.6\pm0.1$ eV and 5.3 ± 0.1 eV with $W=1.5$ eV, as computed; (b) Spicer's results for $P(E)$, $W=1.5$ eV.

with the experimental results obtained by Spicer[48] and by Gobeli and Allen[49] as shown in Fig. 19. The value of ω is experimentally varied, as we said, by covering the Si surface with a fraction of a monolayer of Cs. The over-all agreement appears satisfactory. The principal features in the $W=2.7$-eV plot are the peaks at 3.4 and 5.3 eV due to $\Gamma_{25'}\rightarrow\Gamma_{15}$ and $L_{3'}\rightarrow L_3$ transitions. The $X_4\rightarrow X_1$ dip at 4.3 eV arises since $E(X_1)-E(\Gamma_{25'})=1.1$ eV $<W$. The secondary peaks at 3.8 and 4.8 eV

[48] W. E. Spicer and R. E. Simon, Phys. Rev. Letters 9, 385 (1962).
[49] G. W. Gobeli and F. G. Allen (to be published).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

have been identified as due to transitions near $\Lambda_3 \rightarrow \Lambda_1$ and points in the volume of the Brillouin zone, respectively. As W is increased the behavior of the curves is in quantitative agreement with the experiments.

In Fig. 20(a) we have drawn the energy distribution $P(E)$ for the emitted electrons with $W=1.5$ eV for the cases $\hbar\omega=3.5$ and 5.3 eV. Since our sample includes considerably fewer points, we take an energy interval of 0.3 eV. The experimental results of Spicer are drawn for comparison in Fig. 20(b). Again our results provide an interpretation of the experiment. The peak at 1.2 eV in the $\hbar\omega=3.5$-eV distribution is centered around the $\Gamma_{25'} \rightarrow \Gamma_{15}$ transition. The peak at 2.4 eV in the $\hbar\omega=5.3$-eV curve arises from $L_{3'} \rightarrow L_3$ transitions.

XII. DISCUSSION OF RESULTS

By starting with a model term scheme for the levels at Γ, X, and L, we have constructed a pseudopotential which has given the bands *throughout* the zone. The results we have found for $\epsilon_2(\omega)$ explain the line shapes which are seen in experiment. As an additional feature we can explain the results of recent studies on photoemission from Si.[50] Although our starting model had errors in it, they were subsequently straightened out by the band calculations. We also found several new critical points. This demonstrates the value of the optical data as a guide to theory. It gives us a way to start our calculation and a way to check it upon completion.

We have seen that the calculation was limited by convergence and sampling noise to errors $\sim$0.05 eV. Further improvements of these quantities can be achieved by increasing the number of plane waves admitted to the expansion, and by taking more points into the zone sample. The next step would require the use of a more accurate potential.

By analyzing the optical properties of Ge and Si using a very simple model we have succeeded in exploring the electronic structure of these materials over an energy range of order 10 eV. This range is an order of magnitude larger than that studied by infrared experiments. The precision of our study enables us not only to explain all the observed structure but also to predict new features, such as M_0 thresholds.

Comparing the results for Ge and Si we find that alterations in the crystal potential are small and lead to much smaller changes in the over-all band structure than might have been supposed from infrared data. This agrees with the ultraviolet experimental results. One may say that in Ge and Si the energy surfaces in k space are determined primarily by the shape of the Brillouin zone, i.e., by symmetry considerations, and by $V(3)$, which measures the strength of the covalent bond.

[50] A further application has allowed us to explain some of the properties of Ge-Si alloys as well as high-pressure effects on the band structure [F. Bassani and D. Brust, Phys. Rev. 131, 1524 (1963)].

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor J. C. Phillips for suggesting this problem and for guidance throughout the work, Dr. F. Bassani for valuable discussions, M. Yoshimine for assistance in checking the program, Professor M. H. Cohen for helpful discussions. He further wishes to thank Dr. O. C. Simpson for his hospitality during the author's stay at the Argonne and to the staff of the IBM 704 for aiding the author whenever possible. He also wishes to acknowledge Dr. H. Ehrenreich and Dr. H. Philipp for communicating their results on the imaginary part of the dielectric constant, Marvin L. Cohen for assistance with the data reduction and useful conversations, and his thesis committee for a critical reading of the manuscript.

APPENDIX

In Fig. 21 we plot $K_{4,5}(\omega_i)$. This should be compared with $\bar{K}_{4,5}(\omega_i)$ shown in Fig. 8(a) of the text, which shows the effect of smoothing.

To do a rigorous statistical analysis of the smoothing process would be exceedingly complex and not particularly rewarding. By a few heuristic arguments we can, nevertheless, estimate the random errors entering the problem. Consider first a one-dimensional situation. Suppose we have in this model a monotonically increasing $E_{ns}(\mathbf{k})$ curve. In Fig. 22(a) imagine that the solid lines divide the energy range into intervals of 0.1 eV, that is, $E_{ns}(\mathbf{k}_{i+1})=E_{ns}(\mathbf{k}_i)+0.1$ eV. Next suppose that a uniform net of sampling points is laid over the line. The result of this sample is used to approximate the true joint density of states. Suppose also that in computing $E_{ns}(\mathbf{k})$ we had convergence noise $\sim$0.05 eV. There are then two sources of error in the computed joint density of states. The first results from the fact that the last sampling point in an interval is not significant. For example, in Fig. 22(b) shifting the boundary at 0.2 eV very slightly to the left will, put the last point in the 0.1–0.2 eV interval into the 0.2–0.3 eV interval. The second source of error arises because our calculation does not locate the boundary energies exactly. The idea is sketched in Fig. 22(c). One should note that the error in the energy eigenvalues is a dis-

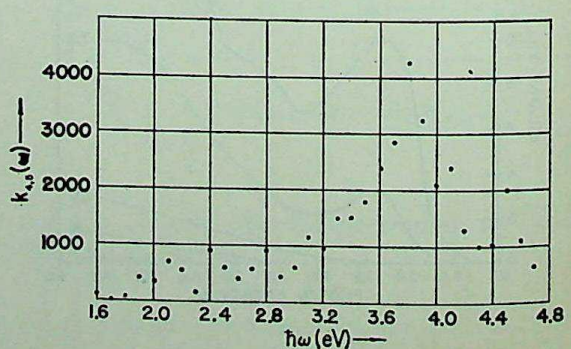


FIG. 21. $K_{4,5}(\omega_i)$, compare with $\bar{K}_{4,5}(\omega)$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

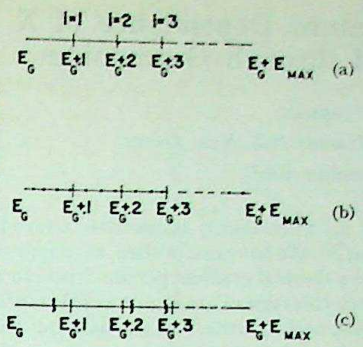


FIG. 22. Monotonically increasing $E_{ns}(\mathbf{k})$ in one dimension (a) energy divided into intervals of 0.1 eV, (b) uniform sampling net used to determine joint density of states, (c) wavy lines indicate erroneous boundaries of intervals arising from convergence noise.

continuous function of $\mathbf{k}$. This occurs because the number of plane waves in our expansion depends on $\mathbf{k}$.

If we take a sampling mesh of 1000 points and suppose that $E_{ns}(\mathbf{k})_{max}=5$ eV, then our one-dimensional sample contains $\sim$20 points per interval. These numbers roughly correspond to those of the actual problem. Then the error from boundary corrections is about 1 point per interval or 5%. The error arising from convergence noise can be $\sim$10 points per interval or 50%. The effect of the smoothing is clear. We take three times as many points while leaving the absolute error constant. Hence the dominant error is reduced from about 50% to about 20%.

The situation in three dimensions is not so readily analyzed. The same sort of arguments ought to apply, however, we now have a two-dimensional surface for the boundaries of the energy intervals. The scatter resulting from boundary corrections and from convergence noise before smoothing is now estimated to be $\sim$20%, and after, $\sim$10%.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Redetermination of the Temperature Dependence of X Rays Anomalously Transmitted through Germanium

Boris W. Batterman
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 2 December 1963)

We have redetermined the temperature dependence of the anomalously transmitted x rays through single crystals of germanium of effective thicknesses, $\mu_0 t = 20$ and 30. We have established, as suggested by Okkerse, that the earlier measurements were subject to error due to a thermal gradient perpendicular to the diffracting planes. By suitable rearrangement of the heating geometry this type of gradient was substantially removed. The current results are in agreement with those of Okkerse and indicate that the data can be explained by allowing the imaginary part of the scattering factor to depend on temperature as $\exp(-M)$ where M is based on a Debye temperature of 291°K. It is felt that there is insufficient evidence at present to conclude in general that the temperature dependence in dynamical diffraction is $\exp(-M)$.

In a previous publication[1] we reported measurements on the temperature dependence of the anomalous transmission of x rays in germanium single crystals. We obtained a linear dependence of the logarithm of the integrated intensity with temperature ranging over an order of magnitude in intensity. [Reference to this work will be abbreviated as (I).] The data in (I) were fitted to theory assuming that $\epsilon(=F_H''/F_0'')$, the normalized angular dependence of the imaginary part of the scattering factor, could be expressed as $\epsilon = \epsilon_0 \exp[-(\alpha M)]$ where ϵ_0 is the static value, M is the exponent in the usual Debye-Waller factor e^{-M} and α is an adjustable constant. It was found that the data for the (220) and (400) reflections were consistent with a single $\alpha = 1.32$.

In a symposium associated with the 50th anniversary celebration of the discovery of x-ray diffraction, Okkerse[2] presented experimental data indicating a similar logarithmic dependence but with an α value of 1.0 based upon a Debye temperature of 291°K.

To resolve the experimental discrepancy, we repeated our past measurements with several variations. In the first series of experiments the crystals were mounted in the same manner as before (crystal in the shape of the letter T, held by the vertical arm) as in (I), Fig. 4, but with the crystal extended in the furnace so that both the forward diffracted (anomalously transmitted) and diffracted beams could be measured. (Previously only the diffracted beam was measured.) The experiment was repeated with the x-ray beam striking different portions of the surface of the crystal. The results of these measurements were: (1) the slopes of log intensity versus temperature of the diffracted beams were constant for different points on the crystal surface to within 5%, but did not agree in value with the previous measurements for a crystal of the same thickness, (2) the curves of the forward diffracted beam were not linear but curved toward lower intensities at higher temperatures such that at sufficiently high temperatures, the forward beam had less intensity than the diffracted beam. The amount of curvature depended upon where the beam struck the crystal surface.

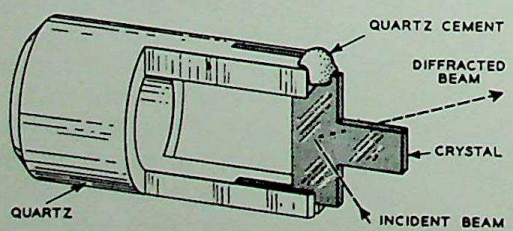


FIG. 1. Details of crystal mount. The crystal is attached on only one arm of the yoke.

These results strongly indicated the presence of thermal gradients even though the slopes of the diffracted beam curves (which were measured in our previous paper) did not vary with beam position on the crystal.

We subsequently repeated our previously reported measurements on the (400) reflection on the same crystal and experimental setup as in (I). The results reproduced those in Fig. 4, Ref. 1, i.e., they showed the curvature when the beam was near the supported part of the crystal, and a linear plot of the same slope as before, when the beam was at the center of the tee-shaped crystal. However, when the beam was moved still further from the support, the plot remained linear, but had a progressively smaller slope the further the beam was from the quartz-supporting rod. This definitely proved that the results in (I) were in error because of a thermal gradient due to heat flow down the supporting rod. This substantiates Okkerse's[2] hypothesis concerning the results in (I) that a linear dependence of the log of intensity versus temperature is not the correct criterion for a strain gradient-free crystal. In this arrangement heat flow down the quartz-support rod produced a gradient perpendicular to the diffracting planes which strongly affected the integrated intensities.

To avoid this difficulty we now mounted the crystal as in Fig. 1. Any gradient caused by heat flow down the supporting rod would now be parallel to the diffracting planes and should have a minimal effect on the integrated intensity. As before, the germanium (220) and

[1] B. W. Batterman, Phys. Rev. 126, 1461 (1962).
[2] B. Okkerse, Philips Res. Rept. 17, 464 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

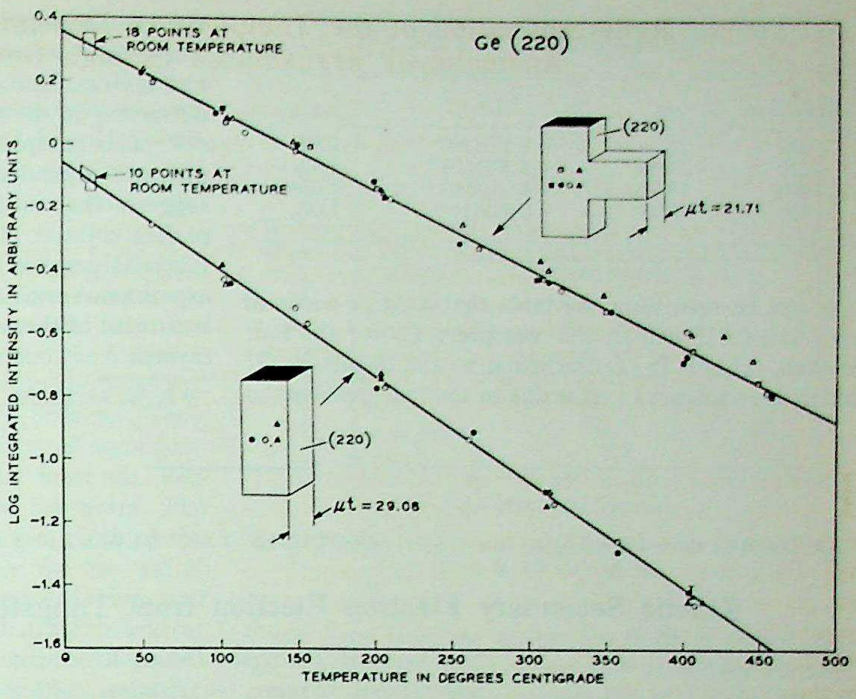


FIG. 2. (220) temperature dependence for two crystals of $\mu_0 t = 21.71$ and 29.08. The insets show the crystal shape and beam position. The shading indicates the mounting face of the crystal.

(400) anomalously diffracted beams were investigated using Cu K_α radiation. Two new specimens oriented for (220) and one new specimen for the (400) were prepared as described in (I). The (400) used in (I) was still available and made the fourth sample. The crystal thicknesses were such that for each reflection we investigated the region of $\mu_0 t \approx 20$ and ≈ 30.

The quantitative measurements were performed in the same manner as in (I), the only variation being the new manner in which the crystals were held (Fig. 1). The data for the two reflections are presented graphically in Figs. 2 and 3. The insets show the positions of the intersection of the primary beam with the crystal surface. This beam had a rectangular cross section of 0.5×0.5 mm^2. The intensity values for the various positions are included in the same plot. The intensities for the different points were taken generally in a random sequence. After measurements at several temperatures for any one crystal position, a room temperature check was made.

The internal consistency of the data for the many beam positions for each crystal gives strong evidence that thermal gradients are not affecting the measurements.

The slope of the curves is related to α of the assumed Debye-Waller factor, $e^{-\alpha M}$, by Eq. (9) in (I) which is repeated below for convenience:

$$\text{slope} = -(1-\alpha/2)k - \mu_0 t \epsilon_0 \alpha k e^{-\alpha M},$$

where $k = M/T$ and M is based on an x-ray Debye temperature for germanium of 291°K.[3] Table I gives the parameters and resulting values for α. The static values, ϵ_0, are very close to unity. For the computation we took Wagenfeld's[4] theoretical values, $\epsilon_0(220) = 0.993$ and $\epsilon_0(400) = 0.985$.

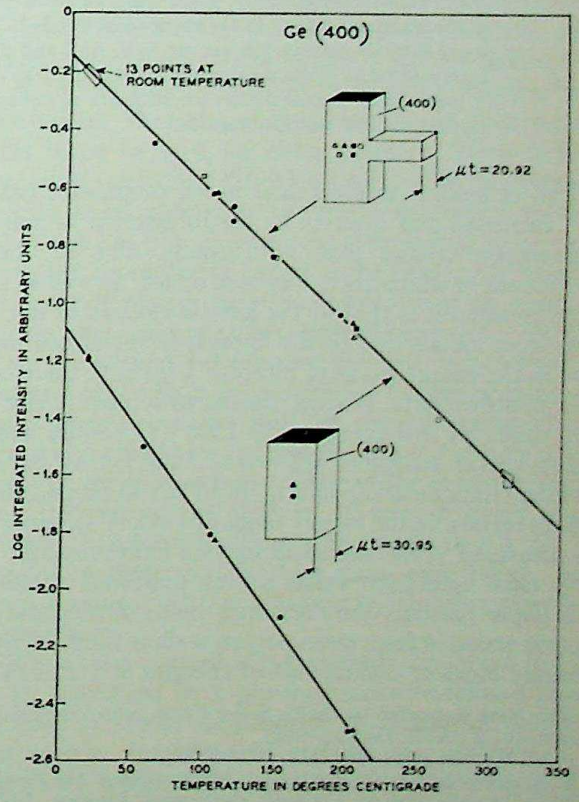


FIG. 3. (400) temperature dependence for two crystals of $\mu_0 t = 20.92$ and 30.95. Details as in Fig. 2.

[3] B. W. Batterman and D. R. Chipman, Phys. Rev. 127, 690 (1962).

[4] H. Wagenfeld, J. Appl. Phys. 30, 2907 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Experimental slopes from Figs. 2 and 3 and corresponding values of α.

Reflection	μt	Slope	α
220	21.71	-2.48×10^{-3}	0.99_9
220	29.08	$-3.39_5\times10^{-3}$	1.02_7
400	20.92	$-4.65_6\times10^{-3}$	1.02_7
400	30.95	-6.85×10^{-3}	1.01_8

It can be seen from the table that a single value of $\alpha=1.01\pm0.015$ will fit the measurements for the four crystals. This is in contradiction to the results in (I) which were subject to error due to thermal gradients in the specimens perpendicular to the diffracting planes. These results now substantially agree with Okkerse's and indicate that Debye-Waller factor for germanium diffracting in the symmetric Laue case is very close to e^{-M}. This, coupled with the same experimental temperature dependence for symmetric Bragg reflection,[5] suggests that in general the Debye-Waller factor for perfect crystals is e^{-M}. This should not, however, be taken as proof of its general validity without further experiments and the development of a sound theoretical treatment of thermal vibrations in an absorbing perfect crystal.

[5] B. W. Batterman, Phys. Rev. 127, 686 (1962).

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Kinetic Secondary Electron Ejection from Tungsten by Cesium Ions*

SAMUEL H. BOSCH AND GUNTIS KUSKEVICS
Electro-Optical Systems, Inc., Pasadena, California
(Received 23 December 1963)

Initial measurements of the number of electrons ejected per incident ion, γ_i, have been made for medium-energy (1–21 keV) cesium ions at normal incidence on a clean polycrystalline tungsten surface. The residual gas pressure was $\leq5\times10^{-9}$ Torr. The tungsten surface was cleaned by prolonged bakeouts and by flashing before each measurement. A dc ion-beam pulse method of measurement was used to prevent cesium coverage of the tungsten surface. In the energy range of 3–21 keV γ_i is a linear function of the ion kinetic energy with $\partial\gamma_i/\partial E\simeq0.04$ electron per ion per keV. At 3 and 21 keV γ_i is 0.02 and 0.74, respectively. If the data are extrapolated to $\gamma_i=0$, a threshold energy of 2.5 keV is obtained. Below this energy the measured values of γ_i were equal to zero within the accuracy of the measurement. The change of γ_i with cesium adsorption on the tungsten was also recorded.

USE of cesium as propellant in ion propulsion has increased the interest in the interaction of medium-energy cesium ions with solids. One of the interactions in which there has been a basic as well as a practical interest is that of the kinetic ejection of electrons from metals due to the bombardment of cesium ions. In the energy range of interest, 1–20 keV, there is little published data. Waters,[1] measured secondary electron yields for low-energy (150–1500 eV) cesium and lithium ions on tungsten. Arifov *et al.*,[2] and Petrov[3] have published data for alkali ions up to 10 keV impinging on various metals. In the energy range of 1–4 keV Brunnee[4] has measured γ_i for the alkali ions on molybdenum.

No close agreement exists among published values. This paper presents the measured yields of secondary electron emission from clean polycrystalline tungsten by normally incident cesium ions of energies of 1–21 keV. Since the ionization energy of cesium, 3.87 eV, is less than the work function of clean tungsten, about 4.5 eV, the electron yields presented are due to the kinetic ejection process only and not to the Auger or potential-ejection[5] process. The apparatus used is shown in Fig. 1.

The vacuum chamber is a stainless-steel cylinder, 3 ft in length with a diameter of 1 ft. It is pumped by a liquid-nitrogen trapped, 6-in. silicon oil diffusion pump. The chamber can be baked to 300°C with either an oven or strap heaters. Copper gaskets are used for all flanges. Within the tank a liquid-nitrogen liner is used to freeze out all condensibles. It can be baked out by passing hot gas through it. Pressure is measured with a Bayard-Alpert type ionization gauge. The system is capable of pressures lower than 1×10^{-9} Torr if proper care and technique are used.

The cesium ions are produced by a surface-ionization ion source using a porous tungsten ionizer operating at 1000–1200°C.[6] This source can generate up to 5 mA of cesium ion current with a maximum current density at the target of approximately 1–2 mA/cm². The fraction of neutral cesium atoms for the currents used in this

* This work supported by the U. S. Air Force, Aero Propulsion Laboratory, WPAFB, Dayton, Ohio.
[1] P. M. Waters, Phys. Rev. 111, 1053 (1958).
[2] U. A. Arifov and R. R. Rakhimov, Transactions of the Ninth All Union Conference on Cathode Electronics, Moscow, 1959 (unpublished), p. 666.
[3] N. N. Petrov, Fiz. Tverd. Tela 2, 949 (1960) [English transl.: Soviet Phys.—Solid State 2, 865 (1960)].
[4] C. Brunnee, Z. Physik 147, 161 (1957).
[5] H. Hagstrum, Phys. Rev. 96, 336 (1954).
[6] G. Kuskevics and B. Thompson, AIAA J. 2, 284 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

experiment was less than 0.01. This represents neutral cesium incidence rates corresponding to monolayer adsorption times of greater than five hours. All measurements were taken with low ion-current densities of approximately 10 μA/cm^2. The singly charged cesium ions, accelerated by the three electrode structures are monoenergetic except for their thermal velocity distribution.

The secondary electron emission detector consists of the target made of 0.003-in.-thick tungsten filament, and a cylinder collector with slits in front and rear. The latter allows higher pumping speed around the filament. The target can be flashed by direct resistance heating. Biased slits are situated in front of the collector. They collimate the beam, prevent the escape of secondary electrons that are ejected toward the front slit, and extract electrons that may be in the ion beam. The edges of the slits are chamfered and the width of the slits becomes progressingly larger from the first slit to the collector. The values of the incident ion and secondary electron currents are obtained by measuring the voltage drop across precision resistors. These are read out on a Westronic two-channel chart recorder. The collector potential is provided by a dc voltage supply in conjunction with a polarity switch.

During the bakeout and pump down of the vacuum system to the 10^{-9} Torr range (approximately 24 h) the tungsten filament is heated continuously at about 1600°K. The procedure used in measuring is a modification of Petrov's dc pulse method.[3] With the ion beam deflected off of the detector the target is flashed for 15 sec at approximately 2300°K with the collector at a negative potential so that it collects any positive ions evaporated from the surface. Three to four seconds after flashing when the target has cooled below thermionic emission temperatures, the collector potential is changed to a +45 V and the chart recorder is started.

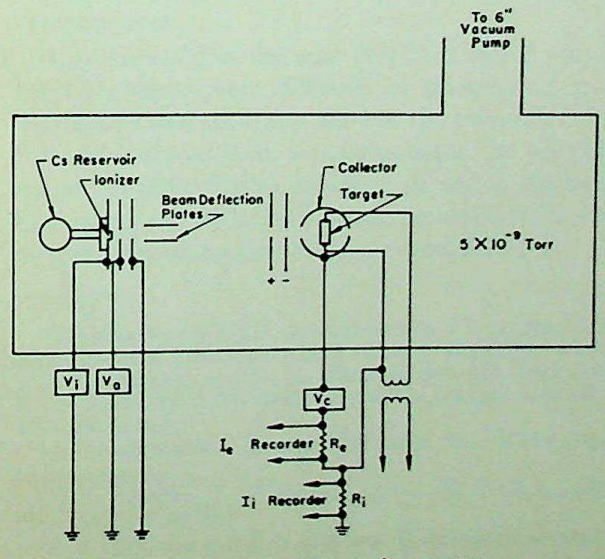


FIG. 1. Schematic diagram of apparatus.

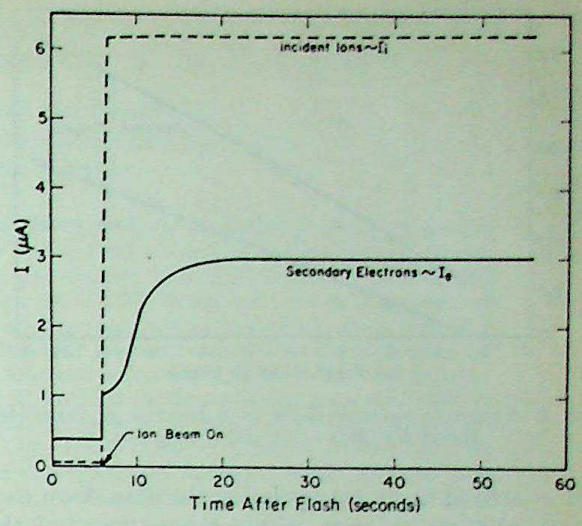


FIG. 2. Typical primary and secondary currents as functions of time. 6 keV Cs^+ on W.

About 2 sec later the cesium-ion beam is pulsed onto the target. The primary ion and secondary electron currents are recorded continuously. If desired, the ion beam can periodically be deflected off of the target to insure the zeros of the ion and electron currents have not shifted.

The instant the beam strikes the target the values of I_i and I_e rise from their zero values to their initial values. These incident ion and secondary electron zero-value currents are due to insulator leakage caused by a cesium film which builds up on them as measurements are taken. The ratio of these initial value increments was taken to be γ_i for clean tungsten. There is little observed variation of I_i with time but I_e immediately begins to increase steadily until a saturation value is obtained (see Fig. 2). Not only is the initial γ_i for clean tungsten reproducible but so is the final γ_i for cesiated tungsten if enough time is allowed. Since all measurements were taken with background pressures $\leq 5\times10^{-9}$ Torr the authors believe that the 5–10 sec delay from the end of the flash to the time of the initial measurement does not allow time for ambient gases to appreciably adsorb on the target. Exact monolayer adsorption times have been measured for these pressures[7] and they are of the order of 10 min. A quick check showed the adsorption time to be greater than 5 min for conditions under which these data were taken. This gives a value of surface coverage by ambient gases less than 3%.

The results of these measurements are shown in Fig. 3. In agreement with previous measurements a linear increase in γ_i was obtained for an increase in incident ion energy, $\partial\gamma_i/\partial E \simeq 0.04$ electron/ion per keV. But contrarily, a higher kinetic ejection threshold energy for cesium on tungsten was obtained. This threshold, 2.5

[7] P. Mahadevan, J. Layton, and D. Medved, Phys. Rev. 129, 79 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

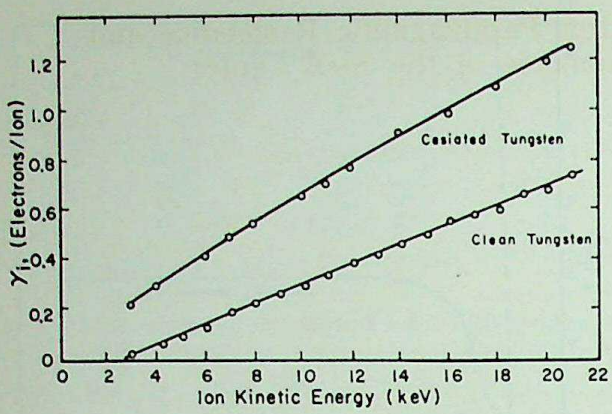


FIG. 3. Secondary electron yields as a function of cesium-ion energy for clean and cesiated tungsten.

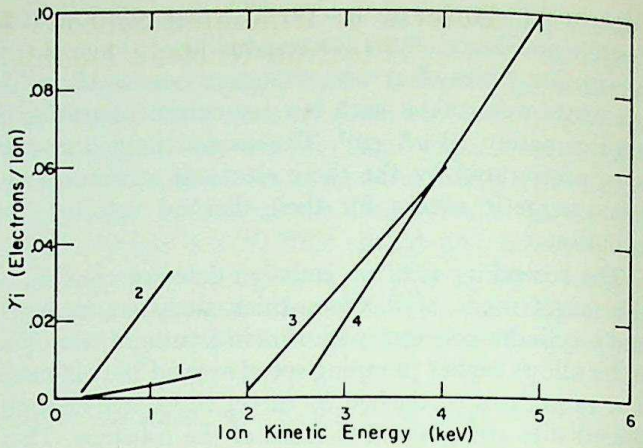


FIG. 4. Comparison of measured values of γ_i as functions of cesium-ion energy. 1—Waters, Cs on W; 2—Petrov, Cs on Mo; 3—Brunnee, Cs on Mo; 4—present data.

keV is arrived at by extrapolating the data down to a $\gamma_i=0$. Also, γ_i was zero within the accuracy of the measurement below this energy for what the authors believe to be clean tungsten. The equipment and method of detection used would be capable of detecting a γ_i down to about 0.005. For cesiated tungsten, nonzero values of γ_i were observed for energies down to 1 keV.

In Fig. 4 the data presented in this paper are shown in comparison with other measurements. Because of the measurement of a threshold energy for γ_i for cesium on tungsten and because of the lower values of γ_i obtained, the authors believe the previous measurements of Waters and Petrov are in error. This error is probably due to cesium or adsorbed gas coverage on the target. In the energy range common to both measurements, the values of Brunnee agree very well with those presented here. There is a difference in the $\partial\gamma_i/\partial E$'s and also in the threshold energy. This could be due in part to the difference in properties of tungsten and molybdenum. The threshold energy of 2.5 keV corresponds to a cesium-ion velocity of approximately 0.6×10^7 cm/sec. This is in excellent agreement with the theoretical value of the velocity threshold for kinetic ejection of electrons calculated by Parilis and Kishinevskii.[8]

ACKNOWLEDGMENTS

The authors wish to thank those whose assistance greatly helped in this experiment. They are Karen Kulseth who assisted in taking and reducing data, Richard Vernon, Stanton Eilenberg, and Frank Barcatta who assisted through discussions on instrumentation and measuring techniques.

[8] E. S. Parilis and L. Kishinevskii, Fiz. Tverd. Tela 3, 1219 (1960) [English transl.: Soviet Phys.—Solid State 3, 885 (1960)]

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Defects in Irradiated Silicon: Electron Paramagnetic Resonance and Electron-Nuclear Double Resonance of the Si-E Center

G. D. Watkins and J. W. Corbett
General Electric Research Laboratory, Schenectady, New York
(Received 7 January 1964)

The Si-E center is one of the dominant defects produced by electron irradiation in phosphorus-doped vacuum floating zone silicon. It introduces an acceptor level at $\sim(E_c-0.4)$eV and gives rise to an electron paramagnetic resonance when this level does not contain an electron. As a result of electron paramagnetic resonance (EPR) and electron-nuclear double resonance (ENDOR) studies described in this paper, we conclude that the defect is a lattice vacancy trapped next to a substitutional phosphorus atom, with EPR arising from the neutral charge state. The observed hyperfine interactions with P^{31} and neighboring Si^{29} nuclei, as well as the observed g-tensor anisotropy, are discussed in terms of a simple linear combination of atomic orbitals (LCAO) molecular orbital treatment. In addition to the anisotropy associated with the phosphorus-vacancy direction in the lattice, an additional distortion of the defect occurs which is identified in the LCAO treatment as a manifestation of the Jahn-Teller effect. Thermally activated reorientation from one Jahn-Teller distortion to another causes motional broadening and narrowing effects upon the EPR spectrum in the temperature region 60–150°K. The motion is also studied by stress-induced alignment at lower temperatures and the activation energy for this process is determined to be $\sim$0.06 eV. Alignment of the phosphorus-vacancy direction in the lattice is also achieved by stressing at elevated temperatures. The activation energy for this motion is 0.93 ± 0.05 eV. The magnitude and sense of the alignment in both kinds of stress experiments are consistent with the microscopic model of the defect. The role of the phosphorus-vacancy interaction in the diffusion of phosphorus in unirradiated silicon is discussed. Using the published value for the diffusion activation energy for phosphorus in silicon, we estimate the appropriate value for silicon self-diffusion to be 3.94 ± 0.33 eV and the formation energy for the lattice vacancy in silicon to be 3.6 ± 0.5 eV. These are quantities for which no direct experimental values are available. Also included is an appendix which gives estimates of $|\psi_{3s}(0)|^2$ and $\langle r_{3p}^{-3}\rangle$ for the $3p$ atoms aluminum through chlorine.

I. INTRODUCTION

THIS is the third[1] in a series giving detailed descriptions of defects produced by irradiation in silicon. In this paper, we describe electron paramagnetic resonance (EPR) studies of the Si-E center.[2] This center has been discussed briefly in previous publications[3,4] where it was tentatively identified as a phosphorus-vacancy pair. In this paper the arguments leading to this identification are presented, and a model of the microscopic configuration of the defect as well as its electronic structure is presented. Such physical properties as its activation energy for diffusion, its binding energy, etc., are estimated.

Also discussed is the role that this defect plays in the high-temperature diffusion of phosphorus in unirradiated material. From some of the properties of the E center deduced here, we can estimate the activation energy for self-diffusion in silicon as well as the formation energy for the lattice vacancy, quantities for which no direct experimental values are available.

II. EXPERIMENTAL PROCEDURE

The silicon samples[5] studied were n type (phosphorus $\approx10^{15}-10^{16}/cm^{-3}$) which had been grown in vacuum by the floating zone technique. The samples were approximately 0.1 in.$\times$0.1 in.$\times$0.5 in. and were irradiated equally on opposite sides at room temperature by 1.5 MeV electrons from a resonant transformer accelerator. The bombardment current was ≈2.5 $\mu A/cm^2$, the maximum temperature rise in the sample being about 25°C.

The samples were mounted with their long dimension along the axis of a TE_{011} microwave cavity and EPR was studied using a balanced bolometer spectrometer at 20 kMc/sec. Magnetic field modulation and lock-in detection at 94 cps were employed. Most measurements were made in dispersion. At $\lesssim$20°K, relaxation times are long and rapid passage effects[6] gave absorption-like spectra, 180 deg out of phase with respect to the modulation. Studies at higher temperatures gave the usual derivative of dispersion curves.

For electron-nuclear double resonance (ENDOR)[7] studies, a single turn coil on the end of a coaxial line was introduced into the resonance cavity, as shown in Fig. 1, the open loop at the bottom of the coil allowing it to slip over the previously mounted sample. The coil was driven by a General Radio 1211B Unit Oscillator

[1] The first two are (I) G. D. Watkins and J. W. Corbett, Phys. Rev. 121, 1001 (1961) and (II) J. W. Corbett, G. D. Watkins, R. S. MacDonald, and R. M. Chrenko, *ibid.* 121, 1015 (1961). In the remainder of this article they will be referred to as (I) and (II), respectively.

[2] In the remainder of the paper, the prefix "Si-" will be dropped for convenience.

[3] G. D. Watkins, J. W. Corbett, and R. M. Walker, J. Appl. Phys. 30, 1198 (1959).

[4] G. D. Watkins and J. W. Corbett, Discussions Faraday Soc. 31, 86 (1961).

[5] Procured from Merck Chemical Company.

[6] A. M. Portis, Phys. Rev. 100, 1219 (1955), Technical Note No. 1, Sarah Mellon Scientific Radiation Laboratory, University of Pittsburgh, November, 1955 (unpublished); M. Weger, Bell System Tech. J. 39, 1013 (1960).

[7] G. Feher, Phys. Rev. 103, 834 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

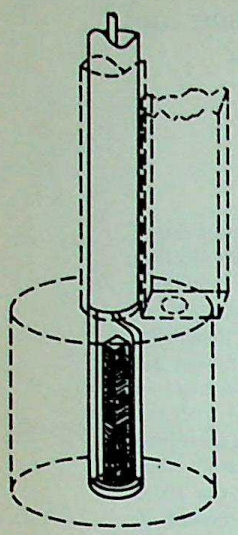


FIG. 1. Single-turn coil assembly introduced into the microwave cavity for ENDOR studies.

and supplied the radio-frequency magnetic field necessary for the nuclear transitions. ENDOR studies were made at 4.2°K by observing a strongly saturated electronic transition in dispersion under normal conditions of magnetic field modulation. The nuclear transition frequencies were detected as a transient change in the electron resonance signal as the radio-frequency oscillator was swept through the nuclear resonance frequency.

The method for applying uniaxial stress to the crystal has been previously described in (I). Briefly, a simple lever and weights outside the cryostat supplies a force to a stainless-steel rod (inserted into the entry tube in place of the ENDOR coaxial line) which in turn transmits this force directly to the end of the crystal, thus compressing it along its long dimension.

Studies versus temperature were made below 77°K by first cooling to 20°K, expelling the liquid H_2 coolant, and observing as the cavity plus crystal warmed up to the 77°K of the cryostat heat shield. For higher temperatures, a heater connected to the wave guide approximately 6 in. above the cavity was used. Temperature was monitored by a copper-constantan thermocouple on the cavity.

III. GENERAL RESULTS AND DISCUSSION

A. Experimental Results

1. Summary of Previous Work

Previous studies[3] have shown the following:

(a) The E center is the dominant spin-resonance center produced in phosphorus-doped *vacuum floating zone* silicon by 1.5 MeV electron irradiation at room temperature. It does not appear in the initial stages of the irradiation but only after the Fermi level has receded to $\approx(E_c-0.4)$eV as a result of the irradiation. The E center is thus associated with a net acceptor level at $\approx(E_c-0.4)$eV[8] and *is observed only when this level does not contain an electron.*

[8] We found this level to be at 0.43 eV. Other investigators have reported the dominant acceptor level in phosphorus-doped floating zone silicon to be at (i) 0.38 eV [G. K. Wertheim and D. E. Buchanan, J. Appl. Phys. **30**, 1232 (1959)], (ii) 0.4 eV [H. Saito, M. Hirata, and T. Horiuchi, J. Phys. Soc. Japan **10**, Suppl. III, 246 (1963)], and (iii) 0.47 eV [E. Sonder and L. C. Templeton, J. Appl. Phys. **34**, 3295 (1963)]. These discrepancies may reflect the presence of other defects with energies near to that of the E center (see Ref. 4).

(b) Hyperfine interaction with 4.7% abundant Si^{29} has shown that the E-center resonance arises from an unpaired electron that is highly localized on one silicon atom. An additional hyperfine interaction with an impurity atom with a 100% abundant nuclear isotope of $I=\frac{1}{2}$ is also observed. This was suggested to be due to a neighboring phosphorus atom.

(c) The E center has not been observed in *pulled* n-type silicon (containing oxygen $\approx10^{18}$ cm^{-3}). Instead in this material the dominant defect observed is the Si-A center,[1] which we have identified as a substitutional oxygen atom, presumably formed when a mobile lattice vacancy is trapped by an interstitial oxygen impurity. This reciprocal behavior between Si-A and Si-E centers, depending upon the concentration of oxygen, led us to suggest that the E center also results from a trapped vacancy, in this case trapped by a phosphorus atom.

2. EPR Spectrum

Figure 2 shows a recording of the E-center spectrum. There is a strong central group made up of doublets and, at higher gain, satellites are observed showing the same doublet structure. Figure 3(a) shows the angular variation of the central group. The satellites are found to be composed of pairs symmetrically located around each of the central lines and with intensity (per pair) $\approx5\%$ of the corresponding central component.

The complete spectrum can be described as arising from several identical defects, but differently oriented in the lattice, each with the spin Hamiltonian

$$\mathcal{H}=\beta\mathbf{H}\cdot\mathbf{g}\cdot\mathbf{S}+\sum_j \mathbf{I}_j\cdot[\mathbf{A}_j\cdot\mathbf{S}-(\mu_j/I_j)\mathbf{H}], \qquad (1)$$

with $S=\frac{1}{2}$.

The second term describes magnetic hyperfine interactions with nuclei neighboring the defect. In these

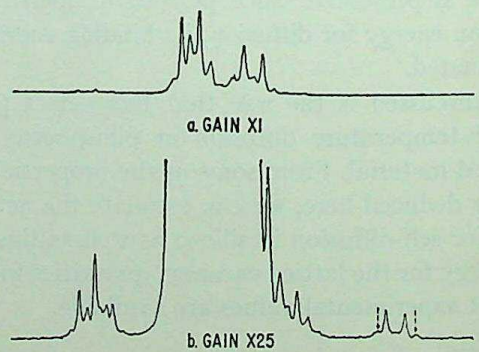


FIG. 2. E-center spectrum at two gain settings with $\mathbf{H}\|\langle111\rangle$. The arrows indicate weak shoulders on the satellite lines, which are referred to in the text.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

interactions, there are two distinct nuclear species involved. One is a nucleus with a 100% abundant isotope of $I=\frac{1}{2}$. (We will demonstrate in the next section that this nucleus is that of a phosphorus atom near the defect, hence we will refer to it here as phosphorus.) There is always one and only one such phosphorus nucleus near the defect, giving rise to the doublet structure on all lines shown in Figs. 2 and 3. The other nuclear species is Si^{29} (4.7% abundant, $I=\frac{1}{2}$) associated with neighboring silicon atoms. The satellites of Fig. 2 arise from strong interaction with the Si^{29} nucleus of *one* nearby silicon atom, their amplitudes relative to the central lines reflecting the isotopic abundance of Si^{29}. Hyperfine interactions with other silicon atoms in the vicinty of the defect are an order of magnitude smaller and are only partially resolved, contributing mainly to the breadth of the central lines.

Equation (1) gives for the energy levels

$$E=g\beta HM+\sum_j E_{hf}{}^j(M,m_j)\,, \qquad (2)$$

with

$$g^2=g_1{}^2n_1{}^2+g_2{}^2n_2{}^2+g_3{}^2n_3{}^2\,, \qquad (3)$$

where n_1, n_2, n_3 are the direction cosines of the principal axes of the g tensor with respect to **H**. Ignoring the small anisotropy in g, the hyperfine energy terms in (2) are given to first order in $A/g\beta H$, $\mu/g\beta I$ by

$$E_{hf}{}^j(M,m_j)=\frac{M}{|M|}m_j\{\sum_\alpha[MA_{j\alpha}-(\mu_j/I_j)H]^2n_{j\alpha}{}^2\}^{1/2}. \qquad (4)$$

Here $n_{j\alpha}$ are the direction cosines of the $\alpha=1,\ 2,\ 3$ principal axes of the $\mathbf{A}_j$ tensor with respect to **H**. Equation (4) has been written in a form convenient for use when $|MA_{j\alpha}|>|\mu_j/I_j|H$, the case for the hyperfine interactions studied here. In so doing, the selection rules for the allowed EPR transitions are the usual $\Delta M=\pm1$, $\Delta m_j=0$,[9] giving

$$g\beta H=h\nu_0-\sum_j[E_{hf}{}^j(\tfrac{1}{2},m_j)-E_{hf}{}^j(-\tfrac{1}{2},m_j)]. \qquad (5)$$

The analysis is summarized in Fig. 4, where the principal values of the **g** and **A** tensors are given along with their principal axes for one of the possible defect orientations.[10] From Fig. 4, it is seen that there must be 24 equivalent orientations available for the defect in the cubic silicon lattice. That is, for each of the twelve possible ⟨110⟩ axes for g_2, there are two equivalent

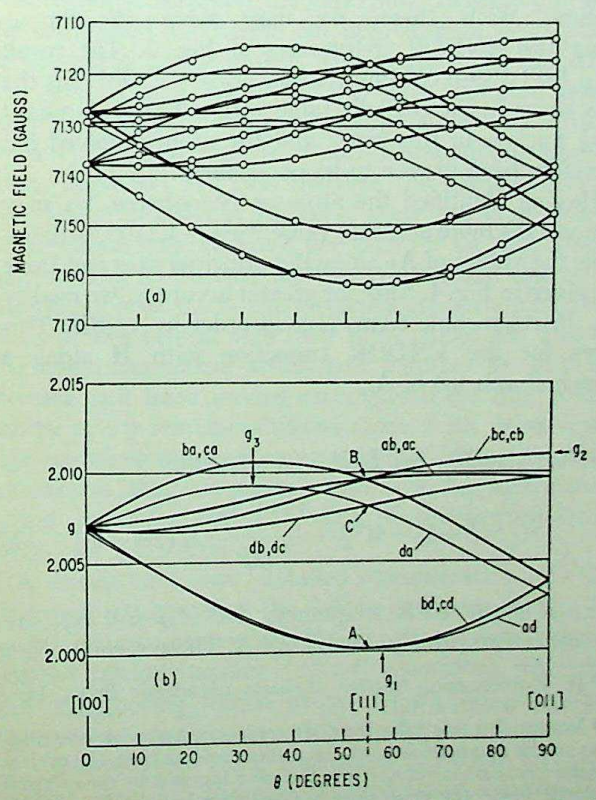


FIG. 3. (a) Angular dependence of the central lines in the E-center spectrum with **H** in the (0$\bar{1}$1) plane (see Fig. 4). The microwave frequency was 20.029 kMc/sec. (b) Corresponding g values. The labeling scheme is described in the text.

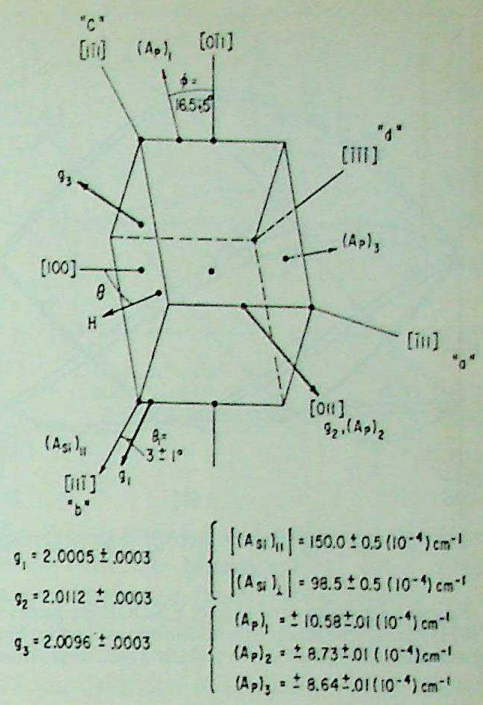


FIG. 4. Spin Hamiltonian constants for the E center with axes indicated for one of the twelve equivalent defect orientations in the lattice. With the magnetic field **H** in the (0$\bar{1}$1) plane as shown, this set of axes gives rise to the g values given by cb in Fig. 3(b).

[9] If $|MA_{j\alpha}|<|\mu_j/I_j|H$, omission of the term $M/|M|$ in Eq. (5) will allow the use of the same selection rules.

[10] These values for A_P were determined also by ENDOR measurements to be described in the next section. They are therefore more accurate than the values determined by analysis of the doublet splitting in the EPR alone. These results are within the stated error of our previously published results (Refs. 3, 4) but reveal in addition the tilting of the $(A_P)_1$ axis by $\varphi=16.5$ deg away from [0$\bar{1}$1], a fact not previously noted.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

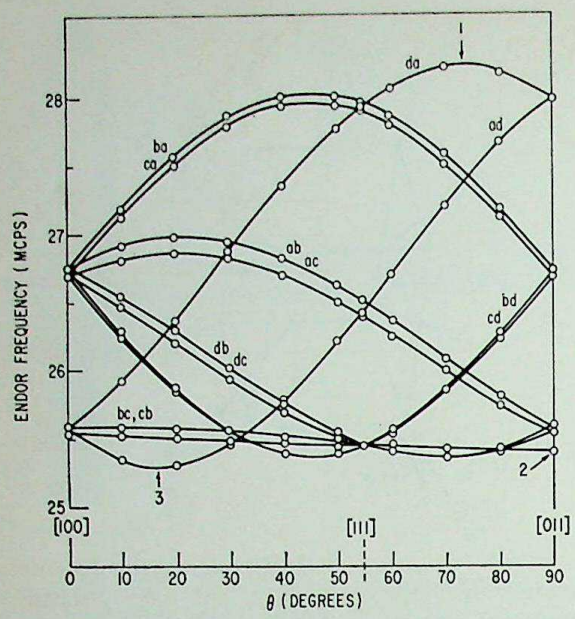


FIG. 5. Angular dependence of the P^{31} ENDOR frequencies with H in the $(0\bar{1}1)$ plane. The labels denote the corresponding branches of the spectrum in Fig. 3(b) for which these transitions are observed.

choices for g_1, one being generated from the other by a 180-deg rotation around g_2. In the EPR spectrum, such a rotation, or inversion, of coordinates cannot be distinguished and the number of distinguishable spectra is reduced to twelve. For most of the studies described here, the spectra were studied with H in a {110} plane (i.e., crystal rotated around the corresponding ⟨110⟩ axis). In this case ten of the spectra become equivalent by pairs, giving only seven distinct spectra as shown in Fig. 3, five of double intensity, and two of unit intensity.

We label (for later use) the transitions of Fig. 3 as referring to a specific defect orientation. This has been done as follows: In Fig. 4, the four ⟨111⟩ axes are labeled "*a*," "*b*," "*c*," and "*d*." (We do not distinguish here the *sign* of the axis direction.) We label a defect by two letters. The first letter specifies the ⟨111⟩ axis closest to $(A_P)_1$ and the second specifies the direction of $(A_{Si})_{II}$.

3. *ENDOR Spectrum*

ENDOR studies were initiated primarily in order to identify the impurity giving rise to the 100% abundant $I=\frac{1}{2}$ hyperfine splitting.[11] From Eq. (4), the nuclear resonance frequencies ($\Delta M=0$, $\Delta m_j=\pm 1$) are

$$\nu_n{}^j(M)=\frac{1}{h}\{\sum_\alpha [MA_{j\alpha}-(\mu_j/I_j)H]^2 n_{j\alpha}{}^2\}^{1/2}. \quad (6)$$

The study was performed only on the higher frequency transition, the lower frequency transition usually being too weak to detect.

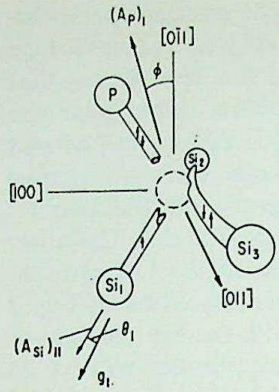


FIG. 6. Simple model of the *E* center as the neutral charge state of a vacancy next to a substitutional phosphorus atom.

The frequencies of the ENDOR transitions versus magnetic field orientation in the {110} plane are shown in Fig. 5 for each of the branches of the spectrum shown in Fig. 3, and are labeled accordingly. Also indicated are the positions of the principal hyperfine axes as determined from these measurements. (The splitting of the *ba–ca*, *bd–cd*, *ab–ac*, *db–dc*, and *bc–cb* pairs arises from ≈1.5 deg misalignment of the ⟨110⟩ axis of the crystal around which the orientation studies were made and would not exist if the crystal were perfectly aligned.)

The $A_{j\alpha}$ can be determined from the doublet splitting in the EPR spectrum, as described in the previous section. With these, we may solve (6) for μ_j using the ENDOR frequencies of Fig. 5. The result is $|\mu_j|=1.13\pm 0.06$. Phosphorus ($\mu_P=+1.1305$[12]) is the only atom in the periodic table with a 100% abundant $I=\frac{1}{2}$ nucleus in this range and the atom involved has therefore been proven to be phosphorus.

Having identified the atom as phosphorus, we may now use the more accurate value,[13] $\mu=+1.1305$, to determine the values of $\mathbf{A}_P$ along the principal axes and these are given in Fig. 4. Also, for greater accuracy, we modify Eq. (6) to include terms second order in $A/g\beta H$. This gives for the ENDOR transition with H along a principal axis (α) of $\mathbf{A}_j$,

$$(h\nu_n{}^j)_\alpha=\Big|A_{j\alpha}M-(\mu_j/I_j)H-(A_{j\beta}A_{j\gamma}/2h\nu) \times[S(S+1)-M^2]-\frac{A_{j\beta}{}^2+A_{j\gamma}{}^2}{4h\nu}M(2m-1)\Big|. \quad (7)$$

Here ν is the EPR frequency, and $A_{j\beta}$ and $A_{j\gamma}$ are the other two principal values of **A**. These second-order

[11] The sensitivity, although adequate for the study of this 100% abundant nucleus, was not sufficient to study the 4.7% abundant Si^{29} hf interactions. As a result this interesting part of the problem, i.e., probing out the wave function on more distant silicon neighbors, has yet to be done.

[12] H. Kopfermann, *Nuclear Moments* (Academic Press, Inc., New York, 1958), p. 451.

[13] In assuming this value for μ, we are ignoring possible contributions to the magnetic field at the nucleus arising from the orbital angular momentum induced into the wave function by the external magnetic field—the equivalent of the "chemical shift" in NMR studies. Using arguments similar to those given by G. W. Ludwig and H. H. Woodbury [Phys. Rev. 117, 1286 (1960)], we conclude, from the small g shifts encountered here, that the effect is negligible with respect to the experimental errors indicated in Fig. 4.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

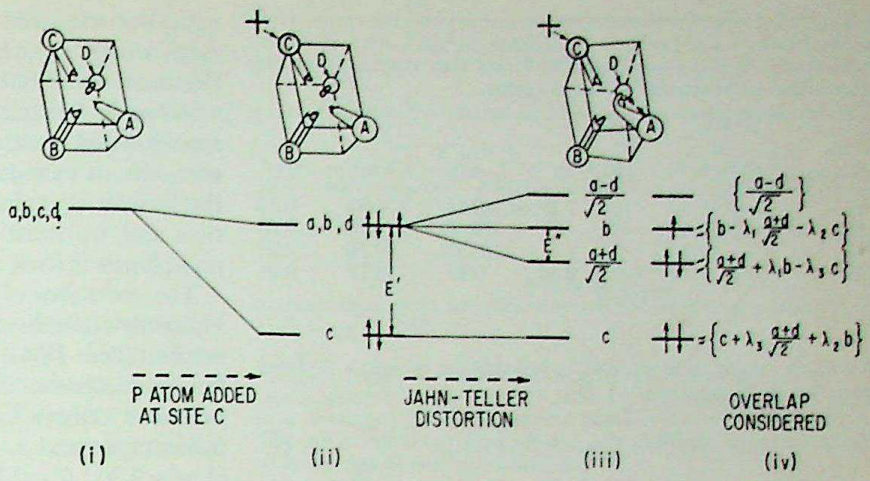


FIG. 7. Simple LCAO molecular orbital model of the electronic structure of the E center. (i) Four degenerate (ignoring overlap) orbitals of the isolated vacancy. (ii) The lowering of the c orbital due to the extra nuclear charge of the phosphorus atom at C. (iii) Lifting of the degeneracy of the a,b,d orbitals by a Jahn-Teller distortion. (iv) Modifications to the wave functions when small overlap between the atomic orbitals is considered. (These wave functions are orthogonal and normalized only through first order in the λ_i.)

corrections are small, and were therefore used only in determining the principal values given in Fig. 4. Equation (6) was otherwise quite satisfactory for fitting the observed spectra. The absolute signs of $A_{j\alpha}$ could not be determined, but fitting the data of Fig. 5 with Eq. (6) requires that the signs of all principal values be the same.

B. Discussion

1. Model

In order to interpret the experimental results, it will be convenient to present an approximate description of the model at the outset. This is shown in Fig. 6 for the defect giving rise to the specific sets of axes shown in Fig. 4. It is the neutral charge state of a lattice vacancy trapped next to a substitutional phosphorus atom impurity. We can visualize the construction of the defect as follows: Initially, there are four broken bonds around the vacancy, one each for the four atoms surrounding it. Because of the extra nuclear charge of the phosphorus atom, two electrons are accommodated in its broken orbital with their spins paired off. Two of the remaining silicon atoms pull together to form an electron pair bond leaving an unpaired electron in the orbital of the remaining silicon atom. EPR arises from this unpaired electron explaining the large hyperfine interaction with one silicon atom. The $\mathbf{A}_{Si}$ and $\mathbf{g}$ axes reflect primarily the $\langle 111\rangle$ axis of the broken bond orbital on this silicon atom.

A somewhat more detailed consideration of the electronic structure of the defect is shown in Fig. 7. In the absence of overlap, the dangling orbitals of an isolated lattice vacancy (a,b,c,d) associated with the four neighboring silicon atoms (A,B,C,D) give four degenerate one-electron orbitals. Replacing atom C by a phosphorus atom is equivalent to adding an extra positive charge at this site. This partially removes the degeneracy and orbital c is lowered in energy. In the neutral charge state, five electrons are accommodated in these orbitals, two paired in the c orbital and the remaining three in the a,b,d set. Because of the degeneracy associated with the partially filled a,b,d orbitals, a spontaneous distortion takes place to remove this degeneracy, as shown. Viewed this way, the "pair bonding" of silicon atoms 2 and 3 in Fig. 6 can be considered simply as a manifestation of the Jahn-Teller[14] effect.

If we now consider the effect of small additional overlap between the remaining orbitals, the wave functions will be modified slightly as shown in panel (iv) of Fig. 7. In this scheme, the unpaired electron is primarily located in orbital b of silicon atom B but has a small percentage ($\lambda_1^2/2$) on silicon atoms A and D and (λ_2^2) on phosphorus atom C. Because we expect the Jahn-Teller splitting (E'') to be smaller than the Coulomb energy associated with the lowering of the c orbital (E'), we expect $\lambda_1^2>\lambda_2^2$.

This is still a considerable oversimplification, of course, because the electronic wave functions will also spill over somewhat onto the more distant neighbors. However, analysis of the hyperfine interactions in the following section shows that $\sim 70\%$ of the wave function can be accounted for on these four atoms alone and this simple localized "molecule" should therefore be a reasonable first approximation.

In the next two sections the hyperfine interactions and the g tensor will be considered quantitatively in terms of this model. In the remaining sections, other consequences of the model will be explored as quantitative tests for the model.

2. Hyperfine Interactions

Following the arguments of the previous section, let us construct the wave function for the unpaired electron as a linear combination of atomic orbitals centered on the atom sites near the defect

$$\bar{\psi}=\sum_i \eta_i\psi_i. \tag{8}$$

[14] H. A. Jahn and E. Teller, Proc. Roy. Soc. (London) **A161**, 220 (1937).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Hyperfine parameters (a_j and b_j) and the corresponding molecular wave function coefficients (α_j^2, β_j^2, and η_j^2) calculated from the observed hyperfine interaction constants for the atom sites in the vicinity of the E center.

Atom site	a_j (10^{-4} cm^{-1})	b_j (10^{-4} cm^{-1})	α_j^2	β_j^2	η_j^2
Si_1	(−)115.7	(−)17.2	0.14	0.86	0.59
Si_2	(−) 12.4	···	(0.30)	(0.70)	0.03
Si_3	(−) 12.4	···	(0.30)	(0.70)	0.03
P	(+) 9.32	(+) 0.63	0.29	0.71	0.01

At each atom site j, we approximate ψ_j as a hybrid $3s3p$ orbital given by

$$\psi_j=\alpha_j(\psi_{3s})_j+\beta_j(\psi_{3p})_j. \tag{9}$$

For the atoms adjacent to the vacancy, we take the p function as directed approximately along the ⟨111⟩ direction from the site to the center of the vacancy. To a good first approximation, the hyperfine interaction at the jth nuclear site is determined solely by ψ_j, i.e., that part of the wave function close to the nucleus. In this approximation, the hyperfine interaction is axially symmetric along the p-orbital axis and can be written

$$A_{\parallel}=a+2b,$$
$$A_{\perp}=a-b. \tag{10}$$

Here the isotropic term a arises from the Fermi contact interaction

$$a_j=(16\pi/3)(\mu_j/I_j)\beta\alpha_j^2\eta_j^2|\psi_{3s}(0)|_j^2, \tag{11}$$

where μ_j is the magnetic moment and I_j the spin of the jth nucleus. The anisotropic term b results from the dipole-dipole interaction averaged over the electronic wave function and is given by

$$b_j=(4/5)(\mu_j/I_j)\beta\beta_j^2\eta_j^2\langle r_{3p}^{-3}\rangle_j. \tag{12}$$

In Appendix A, we have estimated $|\psi_{3s}(0)|^2$ and $\langle r_{3p}^{-3}\rangle$ for neutral silicon and phosphorus, using the recent Hartree-Fock calculations of Freeman and Watson.[15] These values are given in Table III in the Appendix.

Table I gives the results of analysis for α_j^2, β_j^2, η_j^2 for each of the atom sites of Fig. 6, using Eqs. (8)–(12), the hyperfine constants given in Fig. 4 and the estimates of Table III. Equations (11) and (12) predict that a_j and b_j should have the same sign as μ_j. With this assumption, there is no ambiguity in determining the a_j and b_j from (10) even though the absolute values of $A_{\parallel}$ and $A_{\perp}$ were not determined. The signs of the a_j's and b_j's in Table I are in parenthesis to indicate that they were not determined experimentally. In treating the phosphorus hyperfine interaction we have ignored the small departure from axial symmetry $[(A_P)_2\neq(A_P)_3]$ and have ignored the additional contribution to b_j that arises from dipole-dipole interaction with the unpaired electron centered mainly on the neighboring Si_1 site. A crude calculation suggests that this interaction contributes $\sim 0.1-0.2$ (10^{-4} cm^{-1}) to the anisotropic term and it should be included in a more sophisticated treatment. It appears small enough, however, that it cannot account by itself for the tilting of the hf axis away from the ⟨111⟩ "normal bond" direction and we must conclude that the orbital on the phosphorus is itself significantly tilted in this direction.

The estimates of a_j for the Si_2 and Si_3 sites comes indirectly from an experiment to be described in a later section (Sec. IVA), and is included here at this time for completeness. Only a_j can be estimated in this experiment and we have had to assume α_j^2 and β_j^2, i.e., percentage s and p, in order to compute η_j^2. The choice of $\alpha^2\approx 0.30$, $\beta^2\approx 0.70$ was made to agree with that observed in the A center,[16] where the unpaired electron is incorporated in a similar two atom bond between next neighbors across a vacancy.

We are thus able to account for ~60% of the wave function as localized in a dangling orbital on a single silicon atom (Si_1 of Fig. 6, atom B of Fig. 7) neighboring the vacancy. The orbital is approximated by ~86% $3p$ character and is directed toward the vacancy. The enhanced p-like character over a 25%s−75%p, (sp^3), tetrahedral orbital[17] could be caused by the relaxing of the Si_1 atom away from the vacancy. Since the orbital on this atom is not involved in bonding to the other atoms across from the vacancy (see Fig. 7), we would expect the Si_1 atom to be pulled away from the vacancy by the strong bonds to its three remaining nearest neighbors (not shown in Figs. 6 and 7). Si_1 and its three nearest neighbors would then tend to approach a planar configuration for which the bonding orbitals would be sp^2 and the dangling orbital pure p.[18]

An additional ~7% of the wave function is accounted for on the three other atoms surrounding the vacancy, ~3% each on the two silicon atoms (Si_2 and Si_3) and a still smaller amount (~1%), as anticipated in the previous section, on the phosphorus. The remaining ~30% of the wave function is presumably spread over more distant silicon neighbors. [The hyperfine interaction with some of these sites is still large enough to give partially resolved structure on the shoulders of the lines in the spectra (see Fig. 2). The Si_2 and Si_3 interactions contribute to these shoulders, but the total intensity of the structure indicates that there must also

[15] R. E. Watson and A. J. Freeman, Phys. Rev. 123, 521 (1961).

[16] Our previous analysis of the A center (Refs. 1,4) with somewhat different estimates of $|\psi_{3s}(0)|^2$ and $\langle r_{3p}^{-3}\rangle$, indicated that, for it, the wave function was 37% s and 60% p on the main two silicon atoms next to the vacancy and with 70% of the total wave function accounted for by them. A reanalysis, using the new estimates of these quantities given in Table III, gives ~30% s and ~70% p, but with essentially the same fraction of the wave function accounted for. Similar changes, in the fraction s and p character only, are required for the other centers analyzed in Ref. 4.

[17] L. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1948), 2nd ed., Chap. 3.

[18] The arguments are developed somewhat more fully in Ref. 4.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

be three or four additional sites for which the hyperfine interaction is of comparable magnitude. These then could account for an additional 10–15% of the wave function leaving only 15–20% to be further spread out. The weaker hyperfine interaction on these more distant sites is unresolved and contributes only to the breadth (~2.5 G between half-maximum points) of the central part of the line.]

3. *g Tensor*

We have determined that the wave function consists primarily of a dangling bond on a single silicon atom. Let us therefore consider the g tensor to be expected for such a single broken-bond orbital. We start with the Hamiltonian for a bound electron in a magnetic field.[19]

$$\mathcal{H}=2\beta\mathbf{H}\cdot\mathbf{S}+\frac{e\hbar}{2mc}\mathbf{S}\cdot\left[\mathbf{E}\times\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)\right]+\frac{e}{2mc}(\mathbf{p}\cdot\mathbf{A}+\mathbf{A}\cdot\mathbf{p})+\frac{e^2}{2mc^2}A^2. \quad (13)$$

The second term gives rise to the spin-orbit interaction where $\mathbf{E}$ is the electric field seen by the electron, $\mathbf{p}$, its momentum and $\mathbf{A}$, the vector potential. The third term gives the interaction between the motion of the electron and the external magnetic field, and the last term is the usual diamagnetic interaction.

We choose a gauge for $\mathbf{A}$ with the atom involved as center, i.e.,

$$\mathbf{A}=\tfrac{1}{2}\mathbf{H}_0\times\mathbf{r}. \quad (14)$$

With this choice, the third term becomes $\beta\mathbf{H}_0\cdot\mathbf{L}$. Equation (13) then leads to the following expression for the g shift, accurate to first order in the spin-orbit interaction,

$$\Delta g_{ij}=-2\sum_n\frac{\langle 0|(\mathbf{V}_{so})_i|n\rangle\langle n|L_j|0\rangle}{E_n-E_0}. \quad (15)$$

where

$$\mathbf{V}_{so}=(\beta/mc)\mathbf{E}\times\mathbf{p}.$$

The sum is over all excited states.

The form of the g tensor to be expected is apparent from Eq. (15). It will be axially symmetric around the direction of the broken bond orbital axis and the g shift along the axis will be zero.[20] There can be a non-zero g shift perpendicular to this axis, however.

We note that the observed g tensor (Fig. 4) is approximately axially symmetric around the 1 axis (which is approximately the bond axis, see Fig. 6), as expected. Also, the g shift along this axis, although not zero, is

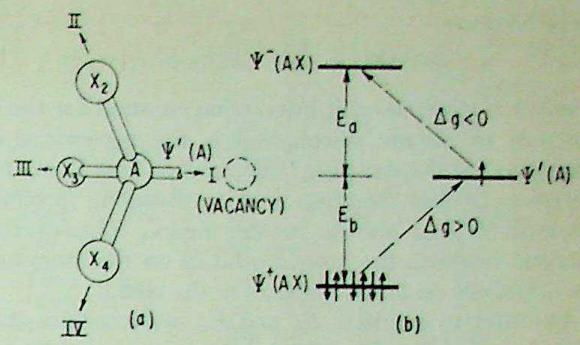


FIG. 8. Localized "molecule" considered in the g-shift calculation (see Appendix B). (a) Atom A with its broken bond $\Psi(A)$ next to a vacancy. (b) Corresponding one-electron molecular states. The matrix elements involved in the g shift are indicated along with the sign of their contribution.

significantly smaller than that along the 2 and 3 axes. If we consider the small negative shift along the bond axis as arising from the 40% of the wave function *not* accounted for in this oversimplified treatment, we are left with $\Delta g_{\parallel}=0$ and $\Delta g_{\perp}\approx+0.01$ to be contributed by the dangling bond. Since $\approx 60\%$ of the wave function is in this bond, this requires $\Delta g_{\parallel}=0$, $\Delta g_{\perp}\approx+0.017$ for a fully occupied dangling bond.

Experimentally, we can take the view that these values are reasonable in that other defect structures in silicon which we have studied and interpreted in terms of dangling bonds also indicate $\Delta g_{\parallel}=0$ and with positive values for $\Delta g_{\perp}$ of this magnitude.[1,4] Theoretically, however, the calculation of these quantities represents a very complex problem. Treated properly, the excited states of Eq. (15) are the continuum of all valence and conduction band states of the solid.

In previous publications,[1,4] we have carried out a simplified treatment of this problem in which we considered only the localized "molecule" made up of the atom in question and its three nearest neighbors as shown in Fig. 8. In this treatment, we used simple one-electron LCAO bonding $[\bar{\psi}^+(AX)]$ and antibonding $[\bar{\psi}^-(AX)]$ molecular orbitals between the atom and its neighbors as approximations to the excited localized valence and conduction band states of (15). The matrix elements to the bonding orbitals, being "hole"-like, were shown to give rise to a positive g shift, while those to the antibonding ones, being "electron"-like, gave rise to a negative shift. The net shift was shown to be the difference between these two *competing* effects. However, it has been pointed out to us that in the derivation, the spin-orbit matrix elements were not handled correctly,[21] leading to an error in the final result. We take this opportunity to correct this error. The result (rederived in Appendix B) should

[19] C. P. Slichter, *Principles of Magnetic Resonance* (Harper and Row, Publishers, Inc., New York, 1963), Chap. 7.

[20] The orbital may be made up of a linear combination of several atomic orbitals (s,p,d, etc.), but by symmetry, all will have $m_l=0$ along the bond axis. As a result, there are no matrix elements for the component of angular momentum along this axis giving $\Delta g_{\parallel}=0$.

[21] We are grateful to R. H. Silsbee for pointing out this error to us. Our error involved replacing $\mathbf{V}_{so}$ by $\lambda\mathbf{L}$, with λ the spin-orbit constant for the free atom. For the molecular excited states considered here, the more general form of (15) should have been used.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

have been

$$\Delta g_1 \approx \lambda\{[1/E_b)-(1/E_a)\}\beta_A^2, \qquad (16)$$

where λ is the spin-orbit interaction constant for the $3p$ electron of atomic silicon, β_A^2 is the percentage $3p$ character for the dangling bond, and E_b and E_a are the energies to the bonding and antibonding orbitals, respectively, as shown in the figure. The overlap integral between the atomic orbitals on the atom and its neighbors no longer appears in the result.

In order to estimate E_b and E_a, we recognize that making up a localized state in the band picture requires summation over all k. E_a and E_b must therefore represent suitable averages over the available states in the bands. Present estimates of the band structure of silicon[22] would suggest that these average excited state energies would be $\sim$1–2 V into the respective bands from the band edges. We do not know the energy-level position associated with the dangling bond of the E center but we *do* know that an *additional* electron can be accommodated at $(E_c-0.4)$eV. This means that the single electron level is lower in energy and, as a guess, we take it as $\approx E_v$, the valence band edge. This gives $E_b \approx 1.5$ eV, $E_a \approx 2.5$ eV. With $\lambda = 0.02$ eV,[23] $\beta_A^2 = 0.86$ (Table I), this gives

$$\Delta g_1 \approx +0.005.$$

This is the correct sign but is a factor of $\sim$3 too low. There was some arbitrariness in the estimates of E_b and E_a, to be sure, but it is not felt that a realistic adjustment of the values could make up this large a factor. (This is particularly evident in the Si-A center,[1] for instance, where the energy level is known to be close to the *conduction band* edge and the g shifts are still positive and of comparable magnitude.)

We thus conclude that the large g shifts must result, in significant part, as a result of stronger matrix elements of (15) to the bonding orbitals (valence band) than to the antibonding ones (conduction band). Within the framework of our localized molecule approximation, there are two effects that we did not previously treat that tend in this direction. One results from the fact that the molecular obritals used in (I) were not orthogonal to the $1s$, $2s$, $2p$ atom cores. We have modified the treatment to take this properly into account (see Appendix B) and we find that (16) is replaced by

$$\Delta g_1 \approx \lambda\{(1+\gamma)/E_b-(1-\gamma)/E_a\}\beta_A^2. \qquad (17)$$

In Appendix B, we estimate $\gamma \sim +0.17$, giving, with $E_b \approx 1.5$ eV, $E_a \approx 2.5$ eV,

$$\Delta g_1 \approx +0.008.$$

This is an increase in the right direction but it is still not enough.

Another effect is a *nuclear shielding* effect. The effect of the positive charge of the neighboring atom core should be to enhance the charge density near the core of the atom to which it is bound for the bonding orbitals but to deplete it for the antibonding ones. This effect is well established in the H_2 molecule, for instance, where the contraction of the $1s$ atomic orbitals in the bonding state is equivalent to an increase of nuclear charge by 20%.[24] Such a contraction in our case would enhance the spin-orbit matrix element of (15) and further increase the "contact" to the valence band.

It would be difficult to include this effect properly in our simple LCAO treatment. However, a measure of its importance is seen in the fact that the observed spin-orbit splitting in the valence band of silicon (0.044 eV)[25] is actually $\sim$50% *greater* than that for the free atom, (0.03 eV).[23] On the other hand, if we were to calculate the spin-orbit splitting using the LCAO molecular bonding orbitals in (I), we would get a *reduction*. For instance, if we ignore the $2p$ core terms, the effective spin-orbit constant would be reduced directly by the normalization factor $(1+S)\cong 1.7$ [see (I)]. By a calculation similar to that for the g shift, we find that inclusion of the $2p$ core terms recovers some of this loss by the factor $1+2\gamma=1.34$, but we would still predict a reduction of $\sim$1.25. This would indicate that the "contraction" of the valence band wave functions is therefore of considerable importance.

It is interesting to note that in our simple treatment, it would require an effective $\gamma \sim 0.8$, to account for the observed 50% increase in valence band spin-orbit interaction. Such a value in (17) would give $\Delta g_1 \approx +0.017$! This agreement is clearly not entirely coincidental in that both the g shift and the enhanced valence band spin-orbit interaction reflect the same phenomenon—the contraction of the wave function around the cores in bonding states.[26]

We conclude that the indicated values $\Delta g_{\parallel}=0$, $\Delta g_1 \approx +0.017$ are *reasonable* ones for a single broken bond orbital in silicon. In the case of the E center, with $\sim$60% of the wave function in such an orbital, this accounts for the gross features of the g tensor, i.e., $\Delta g_1 \approx 0$, $\Delta g_2 \approx \Delta g_2 + 0.01$. The remainder of the wave

[22] J. C. Phillips, Phys. Rev. 125, 1931 (1962).

[23] The spin-orbit interaction constant λ is two-thirds the free atom spin-orbit splitting as given in C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Std. (U. S.) Circular No. 467 (U. S. Government Printing Office, Washington, D. C., 1949).

[24] See C. A. Coulson, Trans. Faraday Soc. 33, 1479 (1937) for a discussion of these effects.

[25] S. Zwerdling, K. J. Button, B. Lax, and L. M. Roth, Phys. Rev. Letters 4, 173 (1960).

[26] The correct characterization of molecular wave functions in the vicinity of the atomic cores is a subject of considerable current controversy [W. Marshall and R. Stuart, Phys. Rev. 123, 2048 (1961); R. G. Shulman and S. Sugano, *ibid.* 130, 506 (1963); J. C. Phillips and L. Liu, Phys. Rev. Letters, 8, 94 (1962)]. In the case of substantial ionic character, the situation is not so clear. In the case here of strong covalent bonds, however, it appears that substantial contraction and expansion of the bonding and antibonding orbitals, respectively, must occur around the atoms. The article by Phillips and Liu, and Ref. 54 treat this point in some detail.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

function (~40%) presumably can account for the remaining g shifts.

IV. MOTIONAL EFFECTS

In the previous sections, we have presented a model for the E center, and we have shown that it is compatible with the observed hyperfine interaction and the g tensor. There is always the question, however, as to whether this model is unique. Are there other possible microscopic configurations that might also be compatible with these quantities? We desire to test the validity of the model in as many ways as possible.

The model of Fig. 6 predicts a number of motional effects which we will test in the following sections. For instance, for the particular phosphorus-vacancy direction shown in the figure, there are two other entirely equivalent configurations, corresponding to bonding Si_1 and Si_3 together and leaving the unpaired electron on Si_2, or Si_1 and Si_2 together, with the unpaired electron on Si_3. These together with the configuration depicted in Fig. 6, represent the three equivalent Jahn-Teller distortions of Fig. 7. The barrier for reorientation from one distortion direction to another should not be large (being comparable to the Jahn-Teller energy) and we expect thermally activated reorientation to manifest itself at relatively low temperatures. This motion can be studied in the resonance in two ways: (1) If the correlation time for the reorientation process is short enough ($\tau \lesssim 10^{-7}$ sec), the individual lines in the EPR spectrum become lifetime broadened and give a direct means of studying the nature and kinetics of the motion. Such a study is described in Sec. IVA; (2) in IVB, the motion is revealed by studying the preferential alignment of the defects induced by the application of uniaxial stress to the crystal. In addition to revealing the nature and kinetics of the reorientation process, the magnitude and sense of the stress induced alignment also confirm other important features of the model.

The vacancy-phosphorus direction is another "degree of freedom" that can be probed by stress. In Sec. IVC, preferential alignment of this axis of the defect is achieved by stress at elevated temperatures. The study of the magnitude and sign of the alignment plus the kinetics of its recovery gives a further test of the model.

A. Low-Temperature Linewidth Studies

In the temperature region ~60–150°K, the E center spectrum is observed to change drastically as shown in Fig. 9.[27] Initially, some of the lines appear to broaden

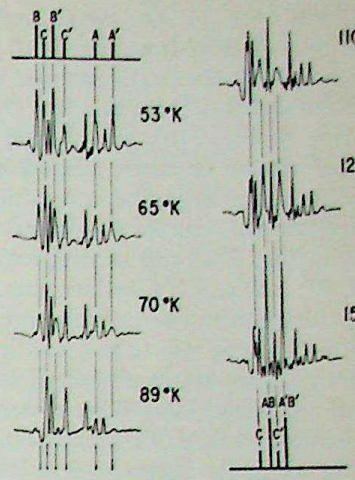


FIG. 9. E-center spectrum versus temperature showing the effect of thermally activated reorientation from one Jahn-Teller distortion direction to another (Ref. 27), $\mathbf{H} \| \langle 111 \rangle$.

and disappear as the temperature is raised and at the higher temperature end a simplified spectrum emerges.

This behavior suggests thermally activated reorientation of the defects. In general, if a defect with an anisotropic g tensor is jumping from one orientation in the lattice to another, its g value will change abruptly with each orientation. In the resonance experiment, this means that, if we are tuned to a particular multiplet in the spectrum (which corresponds to the g value of a particular orientation for the defect), a given defect is alternately brought into resonance, then out, as it executes this random reorientation. This will give rise to a lifetime broadening of the multiplets. As the temperature is raised and the jumping rate increases, the multiplets will tend to broaden out and disappear. At higher temperatures and higher jump rates, motional narrowing will eventually begin and new sets of lines will emerge at the average position of the original multiplets.

Now let us consider the E center in detail and, in particular, determine what we expect if the reorientation motion is from one Jahn-Teller distortion to another. Figure 3(b) shows the variation of the g value versus crystalline orientation for the differently oriented defects. By reference to Figs. 4 and 6, we recognize that the labels on the various curves of Fig. 3(b) have a simple interpretation in terms of the model for the defect. Labeling the four different $\langle 111 \rangle$ axes a through d as in Fig. 4, the first letter of the label specifies the vacancy-phosphorus direction and the second letter specifies the direction from the vacancy to the silicon atom with the unpaired electron. Stated more simply, *for the nearest neighbor sites of the vacancy depicted in Fig. 7, the first letter of the designation gives the phosphorus site and the second letter gives the silicon site for the unpaired electron.* (The defect orientation of Fig. 6 is "*cb*.")

Now consider the thermally activated reorientation from one Jahn-Teller distortion direction to another. In terms of the defect labeling scheme, each time a defect executes a reorientation, its second letter changes,

[27] Also noticed in Fig. 9. are weaker lines, not associated with the E center, which do not change throughout this temperature region. Some of these arise from another center which appears to have roughly the same g variations as the E center and which also displays a phosphorous hyperfine splitting, but of approximately half the magnitude. The interference with the E-center spectrum, and its weaker intensity, has prevented a detailed analysis. We have no model for this defect but presumably it could be an E center paired off with another defect such as oxygen or another vacancy, etc.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

but its first does not. The corresponding g-value change can be read directly from Fig. 3(b).

We now treat the particular case when the magnetic field is parallel to a ⟨111⟩ axis (the conditions of Fig. 9). Referring to Fig. 3(b), the number of distinguishable g values reduces to three, labeled A, B, and C. With the phosphorus hyperfine interaction, this gives rise to three *pairs* of lines for the central group spectrum as shown in Fig. 3(a) and as designated in Fig. 9 as A, A', B, B', and C, C'. Consider now that the magnetic field is tuned to resonance for multiplet C (or C'). For the defects giving rise to this multiplet, the particular reorientation we are considering has no effect because each of the three Jahn-Teller distortions gives rise to the identical g value. (This multiplet arises from the defects for which the magnetic field is pointing along the phosphorus-vacancy direction. By symmetry, it is apparent that each of the three distortions is equivalent with respect to this axis.) Therefore the motion predicts no effect on the CC' multiplets. (There will be a slight narrowing due to the partial motional averaging of Si^{29} hf broadening by neighbors, but we ignore this in this discussion.) In Fig. 9, it is noted that these multiplets (C' is easiest to observe) do indeed stay sharp throughout this temperature range in agreement with the predictions.

The defects with the other three phosphorus-vacancy directions give rise to resonance at $A(A')$ or $B(B')$ and, for them, reorientation causes a jumping between the two resonance values. Gutowsky and Saika[28] have treated a problem which is formally similar to this and we can apply their results directly to this problem. Their results are shown schematically in Fig. 10, in which we consider hopping between two resonance fields A and B separated by $\delta\omega=(g\beta/\hbar)\Delta H$, and with relative probabilities p_A and p_B of being at each of these two fields. (For our case $p_A=\frac{1}{3}$, $p_B=\frac{2}{3}$.) At low temperatures, we have two lines (I) of relative amplitudes p_A and p_B and with widths determined by static interactions (in our case, Si^{29} hf interactions with distant neighbors). As the characteristic reorientation time τ becomes shorter (II), the motion contributes a Lorentzian broadening

$$g(\omega)=\frac{(T_2/\pi)}{1+(\omega_0-\omega)^2T_2^2} \tag{18}$$

to each of the multiplets, where T_2 for each multiplet is given in the figure. At higher temperatures where $\tau\delta\omega\ll 1$, (III), motional narrowing is manifested by the emergence of a line at the weighted average between A and B, which is again Lorentzian broadened with the $(T_2)_{AB}$ given in the figure. As τ becomes shorter the central line narrows further ultimately going over into a temperature-independent width now determined by static interactions that have not been averaged out.

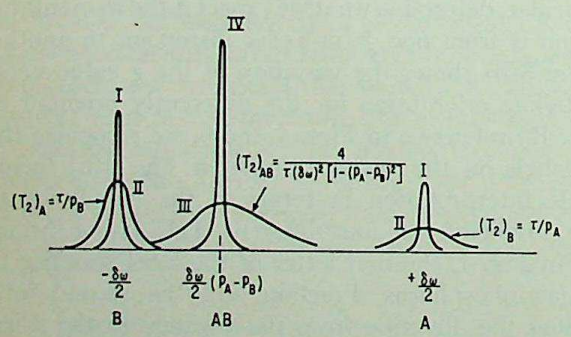


FIG. 10. Effect of thermally activated "hopping" between two resonance positions A and B, with relative probabilities p_A and p_B of being at each site, according to Gutowsky and Saika (Ref. 28). As the temperature increases and the lifetime τ for the "hopping" process decreases, the spectrum changes from (I) low-temperature multiplets at A and B, to (II) broadening of multiplets, to (III) disappearance of multiplets and emergence of a line at the average position AB, to (IV) narrowing of the averaged line.

[28] H. S. Gutowsky and A. Saika, J. Chem. Phys. 21, 1688 (1953). The formulas in Fig. 10 are derived from Eq. (18) of this reference.

Inspection of Fig. 9 shows that the AA' and BB' multiplets broaden as predicted (A' and B show clearly the broadening and final disappearance in the first panel with A' twice the width of B as expected throughout), along with the ultimate emergence and narrowing of the $AB-A'B'$ pair at the weighted average position in the second panel. Similar studies with $\mathbf{H}\|\langle 100\rangle$ and $\mathbf{H}\|\langle 110\rangle$, also confirm the expectations of the model. The motionally averaged spectrum at elevated temperature is axially symmetric reflecting the phosphorus-vacancy direction around which the averaging has taken place. The g values within experimental accuracy agree with those predicted from the weighted averages of the values in Fig. 3(b).

Using the formulas for T_2 in Fig. 10, we can estimate τ from a study of the linewidths versus temperature in both the early broadening and final narrowing regions. The results are given in Fig. 11. {In measurement, the linewidths $\Delta\omega=(g\beta/\hbar)\Delta H$ were taken as the width between half-maximum points observed in the derivative of dispersion recordings as in Fig. 9. For a line of Lorentzian shape, $T_2=1/\Delta\omega$. The points plotted at each temperature represent the *average* of two different methods for subtracting out the contribution due to the static width, $(\Delta\omega)_0$. One method assumed $T_2=[\Delta\omega-(\Delta\omega)_0]^{-1}$, the result expected for the convolution of two Lorentzian shapes. The other took $T_2=[\Delta\omega^2-(\Delta\omega)_0^2]^{-1/2}$, somewhat characteristic of two Gaussians. Since we anticipate the correct behavior to be somewhere in between these two simple approaches, we have simply averaged the two estimates. The accuracy of the data (limited by overlapping of lines, presence of unrelated spectra, temperature uncertainties, etc.) does not warrant a more detailed treatment. We recognize also that a detailed treatment is not trivial. For instance, the static contributions are not strictly constant, being themselves presumably partially motionally narrowed throughout the same region.}

In spite of the uncertainties, we could make a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

reasonably good estimate of τ because the broadening and narrowing regions are separated by about two decades. However, we will defer this until the next section, where additional experimental data allow a more accurate estimate. We note only that the activation energy is $\sim$0.06 eV.

Study of motional effects on the Si^{29} satellites is also informative. This is shown in Fig. 12 for $H\|\langle 100\rangle$. At 20°K, (observed under fast passage conditions in dispersion), we see the Si^{29} satellites with a total splitting of 126 G. In the motionally averaged state (displayed in the figure as derivative of absorption, explaining the apparent difference in the character of the resonance), the Si^{29} satellites are found to be $\approx$ three times as intense relative to the central group but with a splitting reduced to 51 G. This is just what we expect because in the motionally averaged state, the defect becomes a *three* silicon center. That is, in the model of Fig. 6, the unpaired electron is hopping between the Si_1, Si_2, and Si_3 sites so fast as to be indistinguishable from a defect in which the electron is in a LCAO molecular orbital spread over all three. The hyperfine satellites now arise when any one of the three equivalent silicon sites is Si^{29}, which is three times as probable as one, and therefore three times as intense. The hyperfine splitting is determined now by

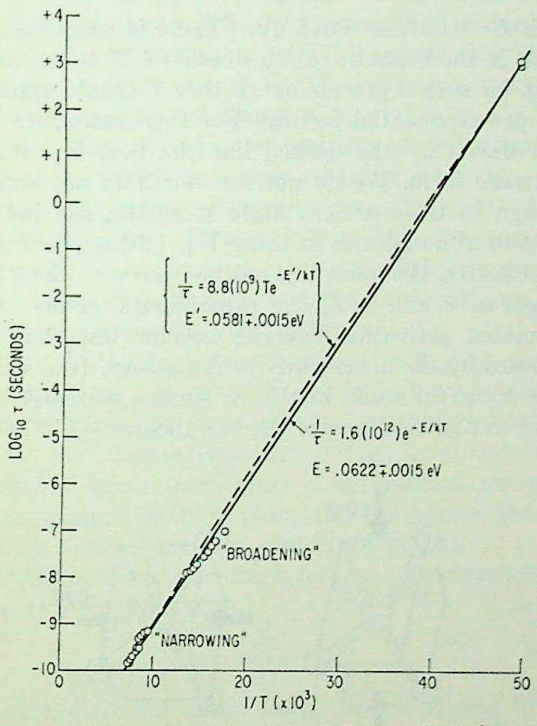


FIG. 11. Lifetime τ of the Jahn-Teller reorientation versus temperature. The two curves represent fits to two different theories (see text).

$$A'(\langle 100\rangle) = \tfrac{1}{3}A_1(\langle 100\rangle) + \tfrac{1}{3}A_2(\langle 100\rangle) + \tfrac{1}{3}A_3(\langle 100\rangle), \quad (19)$$

where $A_1(\langle 100\rangle)$ is the hf term for the defect configuration of Fig. 6 when Si_1 is Si^{29}, and A_2 and A_3 are the correspondingly smaller terms when Si_2 and Si_3 are Si^{29}. The meaning of Eq. (19) is that if one of the three silicon sites is a Si^{29} there is an interaction of A_1 for the third of the time the electron is located mostly on that atom but there is also A_2 and A_3 for each of the other one-thirds when the electron is primarily located on the other two silicon atoms.

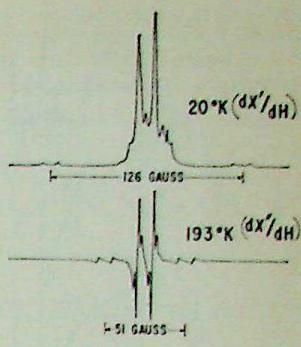


FIG. 12. Comparison of the motionally averaged E-center spectrum at 193°K with that at 20.4°K at $H\|\langle 100\rangle$.

With $|A'(\langle 100\rangle)| = 51$ G, $|A_1(\langle 100\rangle)| = 126$ G, we can solve Eq. (19) for A_2 and A_3, which must be equal by symmetry. This gives

$$A_{2,3}(\langle 100\rangle) = \pm 13.2 \text{ G} = \pm 12.4(10^{-4})\text{cm}^{-1},$$

where the sign must be the same as $A_1(\langle 100\rangle)$. Assuming the interaction to be axially symmetric along a $\langle 111\rangle$ direction, this gives the isotropic constants a_2 and a_3 which are entered in Table I.

Summary: Forgetting the model for a moment, we have determined the following about the center as a result of the linewidth studies: The center undergoes thermally activated reorientation around a particular $\langle 111\rangle$ axis associated with the defect. The activation energy is low ($\sim$0.06 eV), and from the hyperfine studies we have determined that the motion is an *electronic* one. That is, this is *not* the reorientation of a molecular unit but is rather an electronic *bond switching* in which the unpaired electron hops between *three* equivalent atoms which are for all purposes fixed in the lattice.

In terms of the model, this is precisely the behavior expected for thermally activated reorientation of the defect from one Jahn-Teller distortion direction to the other. The 0.06 eV we feel is a reasonably value for the activation energy of the process.[29] This must be considered an important confirmation of the model.

B. Low-Temperature Stress Studies

All of the possible orientations of the E center are equally probable in the cubic silicon lattice and, at

[29] In a simple model, the three silicon atoms go through the symmetrical undistorted position in reorienting. In this case, the 0.06 eV would be equal to the Jahn-Teller stabilization energy. Remembering that this energy is the difference between the total electronic stabilization energy and the increased energy stored in the lattice, which in turn is half the electronic energy, we get for the electronic stabilization energy 0.12 eV. In Fig. 8, this energy is associated with two electrons in the $(a+d)$ orbital so that the energy per electron $E'' \approx 0.06$ eV.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

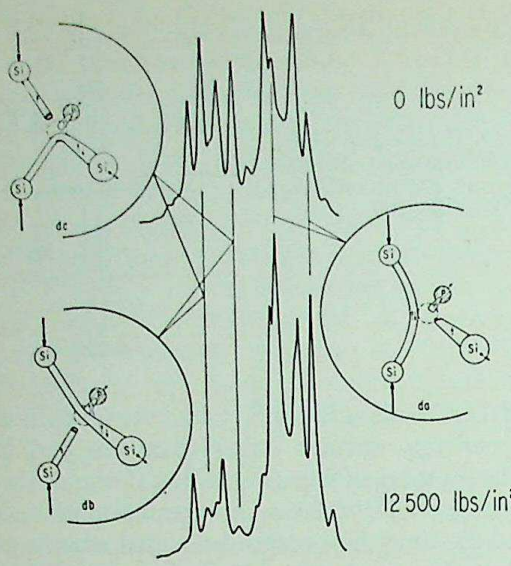


FIG. 13. Change in the E-center spectrum (H‖[011]), under [0$\bar{1}$1] compressional stress. The insets show a typical defect orientation for each multiplet.

thermodynamical equilibrium, we have equal numbers of defects in each of these orientations. If we now strain the crystal, and distort it from cubic symmetry, the various orientations are no longer equivalent and, if the defects are free to reorient, they will do so, tending to seek out the lowest energy orientation. This stress induced alignment is the static analog of the internal friction experiment.

EPR gives a unique way of observing the alignment in that the amplitudes of the multiplets in the spectrum give a direct measure of the relative number of defects in each of the allowed orientations, allowing a microscopic study of the process.[30] We here apply this to study the same "bond-switching" motion that was described in the previous section.

In Fig. 13, the effects on the spectrum of applying a uniaxial compressional stress along a ⟨110⟩ axis of the crystal at 20.4°K is observed. (The spectrum is observed with the magnetic field along another ⟨110⟩ axis, perpendicular to the stress direction.) The insets identify a representative set of defects—and their corresponding letter designation—to be associated with each multiplet. It is seen that there is a large effect, the *dc* and *db* orientations being strongly suppressed and the *da* being proportionally enhanced. This is again consistent with the "bond-switching" motion, the phosphorus-vacancy direction remaining unchanged. (The intensity simply "sloshes" between the indicated multiplets with no change in total area.) The *sense* of the alignment is an *additional* confirmation of the model in that the defect prefers to align in such a way that the applied compression serves to push the two bonding silicon atoms closer together as expected. In terms of Fig. 7, the applied stress adds to the Jahn-Teller distortion and lowers the electronic energy of the defect. We will return to the magnitude of the alignment but first let us consider the kinetics.

Upon removing the stress at 20.4°K, the recovery to the original unstressed spectrum is observed to follow a single exponential law with a time constant $\tau=1.3\pm0.2(10^3)$ sec. This τ is the same characteristic reorientation time that appears in the formulas of Fig. 10 for the linewidth studies and is therefore also plotted in Fig. 11. This value together with the linewidth study estimates span 13 decades in τ, allowing a very accurate determination of τ. The points fall well on a straight line giving

$$\tau^{-1}=1.6(10^{12})e^{-E/kT}, \tag{20}$$

with $E=0.0622\pm0.0015$ eV. [There is some disagreement in the literature as to whether (20) is the correct form for such a process or whether T should appear in the pre-exponential factor.[31] For this reason, we have also shown by the dashed line the best fit with the alternate form. We do not feel our data are accurate enough in their present state to distinguish between the two although the fit using Eq. (20) appears somewhat better. We point this out because with these large ranges of τ and $1/T$, our experimental errors in the estimated activation energies become less than that imposed by the uncertainty in the theory. It is possible that a careful study in this or similar systems[32] could serve as a critical test of the two theories.]

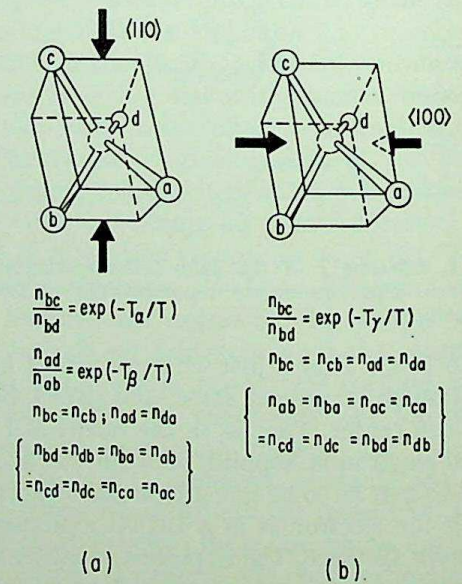


FIG. 14. Boltzmann factors governing the populations of the different defect orientations for the (a) ⟨110⟩ and (b) ⟨100⟩ stress experiments. The subscripts *ij* refer to the defect orientation according to the convention established in the text.

[30] A detailed study of another EPR center (the Si-*A* center) using this technique has been previously described by us in (I). See also W. Kanzig, Phys. Chem. Solids 23, 479 (1962).

[31] See R. H. Doremus, J. Chem. Phys. 34, 2186 (1961) for a discussion of this point and for further references to the literature.

[32] The silicon divacancy [G. D. Watkins and J. W. Corbett (to be published)] also has two spectra which are suitable for such a study.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A quantitative study of the *degree* of alignment has also been performed versus stress and temperature. For this, the effects of stress along both ⟨100⟩ and ⟨110⟩ axes were investigated. By symmetry, the ⟨100⟩ stress creates two nonequivalent sets of orientations while the ⟨110⟩ stress creates three (see Fig. 14). Assuming a Boltzmann distribution between the sets, this defines a single energy difference kT_γ for the ⟨100⟩ stress and two, kT_α and kT_β, for the ⟨110⟩ stress as indicated in the figure.

Fit to a Boltzmann distribution given by the formulas in Fig. 14, the alignment in all cases shows T_α, T_β, and T_γ linearly proportional to stress within experimental accuracy. The exponential dependence upon $1/T$ has also been verified. For this study, estimates in the range 20–50°K were made from the relative intensities of the multiplets as described above. An estimate was also made at 160°K from the *shift* in the resonance field of the motionally averaged spectrum (reflecting the relative change of p_A and p_B in Fig. 10). The fact that *shifts* in the *positions* of the resonance multiplets at high temperatures occur, and that they can be accurately predicted by measurements on the relative *amplitudes* of the multiplets in the low-temperature spectrum confirms again that the high-temperature phase is properly characterized as a motional averaging of the low-temperature configurations.

In Table II, the measured values for T_α, T_β, and T_γ determined at a stress of 12 360 lb/in.^{2} are given. To interpret these, let us take a very simple view of the defect. We assume that the change in energy of a particular defect orientation is determined solely by the component of strain along the line joining the two bonding silicon atoms (Si_2 and Si_3 of Fig. 6).

We define a change in the defect energy per unit strain along this direction as

$$M=(dE/d\epsilon)_{\mathrm{Si-Si}}. \tag{21}$$

Defining ϵ_{ad} as the strain along the line joining the *ad* sites, etc., we obtain

$$\begin{aligned} kT_\alpha&=(\epsilon_{ad}-\epsilon_{ac})M\,,\\ kT_\beta&=(\epsilon_{bc}-\epsilon_{cd})M\,,\\ kT_\gamma&=(\epsilon_{ad}-\epsilon_{ac})M\,. \end{aligned} \tag{22}$$

These follow directly from the formulas of Fig. 14 when the change in energy of the *bc* site is set equal to $M\epsilon_{ad}$, etc. In (I) these strains have been calculated in terms of the elastic moduli and applied stress and, using these formulas with Eq. (22), we can solve for M using the experimental values of Table II. Selection of M for the best fit gives

$$M=+16.8 \text{ eV/unit strain.}^{33} \tag{23}$$

[33] In footnote 29, we concluded that the electronic stabilization energy associated with the Jahn-Teller distortion was ~0.12 eV. This, with Eq. (23) serves to define an "effective" strain associated with the Jahn-Teller distortion of ~0.7%. Whether this in turn bears any simple relation to the actual motions of the atoms neighboring the vacancy is not clear, however.

TABLE II. Comparison of the observed and calculated (see text) energy differences (in °K) for the differently oriented defects under stress (see Fig. 14). The applied stress is 12 360 psi.

Stress	T_j	Observed (°K)	Calculated (°K)
⟨110⟩	T_α	+9	+11
	T_β	−90	−94
⟨100⟩	T_γ	+85	+81

The values computed from Eq. (22) using this constant are also given in Table II.

The agreement in Table II is quite good considering the oversimplified model that we have taken in the analysis. The positive value of M means that the energy is lowered when the atoms are pushed together as we would expect for two normally next nearest neighbors with rather weaker overlap. The magnitude of M is close to the value +16.0 eV/unit strain deduced in (I) for a similar Si-Si pair bond in the A center. As in the A center, we argue that the magnitude is not unreasonable being of the same order of magnitude as deformation potentials in silicon.[34] We thus conclude that the stress experiments give substantial additional confirmation to the model.

C. High-Temperature Stress Studies

There is another "degree of freedom" for the defect, the phosphorus-vacancy axis, which has not been probed as yet. In the low-temperature studies this axis remained fixed. In terms of the model, it should be more difficult for this axis to change because *atoms* must rearrange rather than the easier *electronic* rearrangement involved in the low-temperature motion.

This phosphorus-vacancy axis reorientation can be revealed by stress alignment at elevated temperatures. For this a crystal was compressed along a ⟨110⟩ axis to ~26 000 psi at room temperature in an external squeezing apparatus for ~1 h, then removed and quickly inserted into the microwave cavity and cooled to 20.4°K for study. A significant alignment was achieved as can be seen in Fig. 15. By reference to Figs. 3(b) and 14, we see that the preferred direction for the phosphorus-vacancy axis is in the plane perpendicular to the stress direction.

From Fig. 15, we estimate the degree of alignment[35]

$$\frac{n_\perp}{n_{||}}=\frac{n_{ad}+n_{ac}+n_{ab}}{n_{bc}+n_{ba}+n_{bd}}\cong 1.75. \tag{24}$$

Analysis of Eq. (24), similar to that in the previous section where we consider only the $M=16.8$ eV/unit

[34] C. Herring and E. Vogt, Phys. Rev. 101, 944 (1956).

[35] We estimate the alignment in Fig. 15 to be 1.5. We increase this value by ~15% to take account of partial recovery during the ~75 sec at room temperature between removal of stress and eventual cooling of the sample (see Fig. 16).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

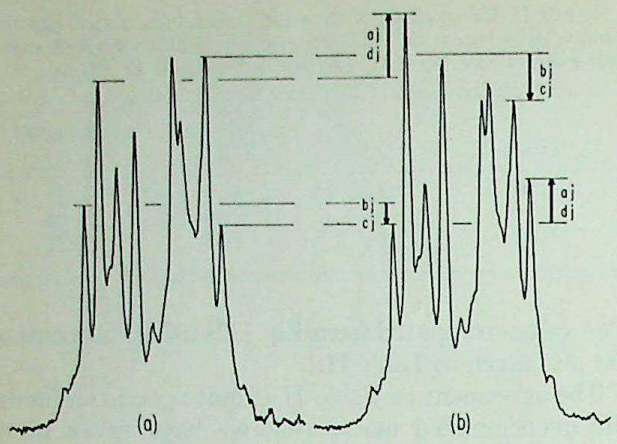


FIG. 15. E-center spectrum (H‖[011]) (a) before and (b) after 26 000 psi [0$\bar{1}$1] compressional stress at room temperature. The defect orientations corresponding to the lines in the spectrum are indicated.

strain associated with the Si–Si bond, predicts $(n_\perp/n_{11})=1.34$. The alignment is therefore in the same sense as that dictated by the preference of the defect to align its Si–Si bond along the compressed direction. If we include another parameter for the change in energy per unit strain at right angles to this, along the phosphorus-vacancy direction,

$$N=(dE/d\epsilon)_{\text{P–Vac}}\cong -9 \text{ eV/unit strain}, \quad (25)$$

the value of 1.75 is obtained. The sign is consistent with an outward relaxation of the phosphorus atom from the vacancy in its normal configuration which is to be expected since there is no bonding to hold it in (see Fig. 7). It is not possible to make a value judgment on the magnitude, but it does not seem out of line with the value for M, for instance.

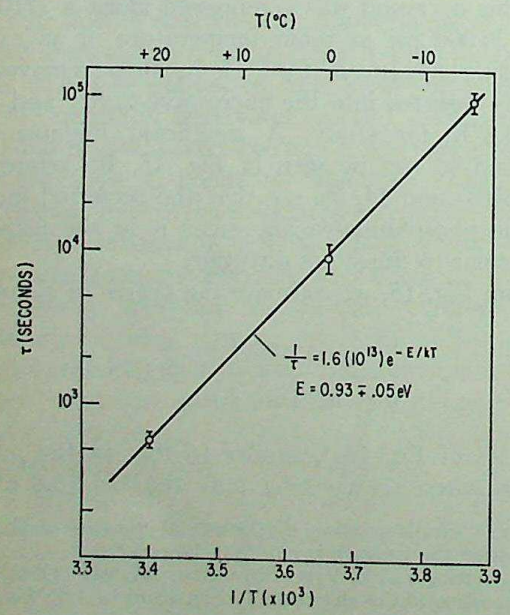


FIG. 16. Characteristic recovery time versus temperature for the phosphorous-vacancy axis reorientation.

The kinetics of the reorientation process was studied by monitoring the recovery from the stress induced alignment versus annealing. During each anneal, the sample remained in the cavity which in turn was held at the annealing temperature for a specified time interval. At the end of each annealing interval, the cavity was plunged back into the cryostat and the remaining alignment monitored from the spectrum at 20.4°K. The results are shown in Fig. 16. We obtain for the characteristic relaxation time of the reorietnation process,

$$\tau^{-1}=1.6(10^{13})\exp(-E/kT),$$

with

$$E=0.93\pm 0.05 \text{ eV}.$$

Let us now inquire as to the microscopic process by which the defect reorients. That is, of what process are we measuring the kinetics? Interchange of the phosphorus and vacancy serves only to "invert" the defect and does not achieve the reorientation we are observing. The most likely remaining process is the one shown in Fig. 17 in which the vacancy must make four consecutive jumps, two away from the phosphorus and then two back. The fact that the defect must partially dissociate in order to reorient has some interesting consequences.

First let us note that the combination of reorientation plus the phosphorus-vacancy interchange allows the defect to diffuse through the lattice as an entity. We anticipate the activation energy for the vacancy-phosphorus interchange to be small, being comparable to the value for the isolated vacancy motion ($\sim\frac{1}{3}$ eV).[36] The limiting process for diffusion should therefore be the reorientation. We thus conclude that in measuring the activation energy for reorientation, *we are also measuring the activation energy for the diffusion of the phosphorus-vacancy pair in the lattice.*

The fact that the vacancy must separate partially from the phosphorus in order for the defect to reorient gives a simple explanation as to why the activation energy is larger than that associated with the isolated vacancy motion. The difference ($\sim$0.6 eV) can be associated with binding energy that must be overcome in order for the vacancy to separate to the next-next-nearest neighbor position (position 3 in Fig. 17). This is not the *total* binding energy, however, because we do not detect any loss in E centers in the process. That is, there is still sufficient binding at this remote position to assure that most of the vacancies return.

There are several bits of experimental information that give us guides in estimating the remaining binding energy at this next-next-nearest site. For instance, we note that the E centers decay upon standing at room temperature at a rate suggestive of a time constant $\lesssim$6 weeks.[37] On the other hand, from Fig. 16, we deter-

[36] G. D. Watkins, J. Phys. Soc. Japan 18, Suppl. II, 22 (1963).

[37] The actual rate may be slower than this. The loss of E centers in the spectrum can also be caused by decay of other centers which in turn alters the Fermi level so that E centers still present are no

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

mine that the defect is reorienting once every ~400 sec at this temperature (~25°C). We may therefore conclude that the vacancy gets away only once in $\lesssim 10^4$ opportunities at the next-next-nearest site. This, in turn, can be related to a Boltzmann factor weighting the separation and recapture processes at this site,

$$\alpha \exp(-\Delta E/kT) \lesssim 10^{-14}, \qquad (26)$$

where α is a geometric factor (~1). Taking $\alpha=1$, we obtain $\Delta E \gtrsim 0.25$ eV. This quantity is not strictly equal to the binding energy at this site but rather gives a difference in barrier heights as shown in Fig. 18. However, it is clearly related (being perhaps somewhat larger) and serves as an indication of its magnitude.

It is interesting to note that a calculation of the binding energy between a single positive and a single negative charge separated by the same distance and immersed in a dielectric constant equal to that of silicon (11.6) gives 0.28 eV. This is certainly suggestive

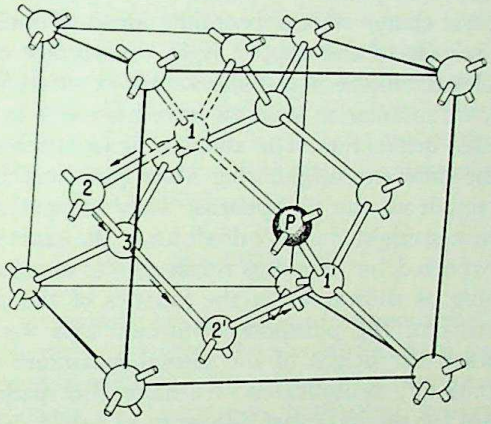


FIG. 17. The mechanism for the phosphorus-vacancy axis reorientation. The vacancy in the normal nearest position (No. 1) must make two jumps away from the phosphorus to site No. 3 and then return to a new (No. 1′) nearest site.

and perhaps not an unreasonable way to view the defect at this separation. The over-all defect is neutral and, in separating into components, the phosphorus will tend to be positively charged, its normal ionized state when isolated.

The indication that $\lesssim 10^4$ reorientations are occurring before disassociation means that the defect is diffusing through the lattice as an entity for considerable distances. Considering the approximate nature of these estimates it is not clear then whether the annealing of the E center actually is a result of its dissociation or whether it is the result of long-range migration as an entity with subsequent trapping by other defects. An experiment that may bear upon this is that reported by Saito *et al.*[38] They have studied by electrical measurements the annealing kinetics of a deep acceptor at $(E_c-0.4)$eV in floating zone silicon produced by irradiation. Considering the conditions of their experiment, it is possible that they are monitoring the E center anneal, as they have suggested. They observe first-order kinetics with a characteristic recovery time τ given by[39]

$$\tau^{-1} \cong 10^{13} \exp(-0.94/kT).$$

The close agreement of this activation energy to that which we estimate for *diffusion* suggests that the annealing they are studying is the diffusion to traps. This is also consistent with the pre-exponential factor of 10^{13} indicating, with our pre-exponential factor of $\sim 10^{13}$ for the single reorientation process, $\sim 10^5$ diffusional jumps before trapping. Analysis similar to that leading to (26), but using $\lesssim 10^5$ jumps without separation and at the higher temperatures of anneal used by Saito *et al.*(~175°C) gives $\Delta E \gtrsim 0.45$ eV. This value is also indicated in Fig. 18.

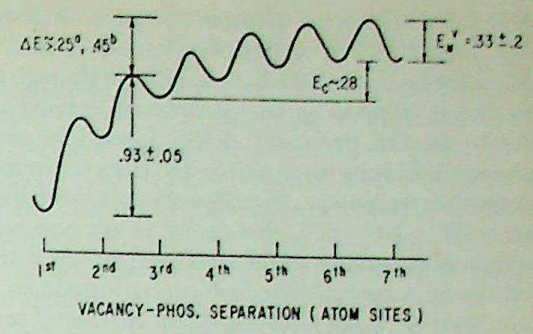


FIG. 18. The potential energy (in eV) of a vacancy in the vicinity of a substitutional phosphorus atom, as determined from the experiments on the E center. The abscissa denotes the various neighbor positions in order of increasing separation, and are not necessarily in a straight line. The quantity ΔE was estimated from annealing studies by (a) EPR and (b) electrical measurements.

On the other hand, there is also evidence for dissociation in that some Si-A centers (oxygen-vacancy pairs) are observed to form as the Si-E centers disappear. This has been observed in electrical measurements[40] and has also been observed by us in resonance. These vacancies could be liberated of course, after trapping. In any event, we recognize from our estimates here that it is touch and go as to whether the vacancy and phosphorus stay together long enough to reach a trap as an entity or whether they dissociate first, and that probably both processes are occurring. To investigate this in detail one would require studies over a wide range of temperatures and defect concentrations.

V. SUMMARY AND CONCLUSIONS

We conclude that the Si-E center resonance arises from the neutral charge state of a lattice vacancy ad-

longer in the proper charge state for resonance. This complication has discouraged us from making detailed annealing studies for this center.

[38] H. Saito, M. Hirata, and T. Horiuchi, J. Phys. Soc. Japan 18, Suppl. III, 246 (1963).

[39] We have estimated the pre-exponential factor from the curves of Fig. 6 of their paper. (See Ref. 38.)

[40] G. K. Wertheim and D. N. E. Buchanan, J. Appl. Phys. 30, 1232 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

jacent to a substitutional phosphorus atom. It is presumably formed when a lattice vacancy, formed at random in the lattice by irradiation, diffuses through the lattice, and is trapped by the phosphorus impurity. It appears to be the dominant defect formed in phosphorus-doped floating zone silicon by electron irradiation at room temperature. In pulled crystals, containing oxygen $\sim 10^{17}-10^{18}$ cm^{-3}, this center is not observed, the oxygen now serving as the dominant vacancy trap to form the Si-A center.[1] The Si-E center introduces an acceptor level at $\sim(E_c-0.4)$eV. With the Fermi level position above this level, the defect is negatively charged and is not observed in the resonance.

The model of the center deduced from the studies in this paper is one in which two of the three silicon atoms adjacent to the vacancy pull together to form a pair bond with the unpaired electron observed in the resonance left primarily to the third silicon atom. This distortion can be viewed as a manifestation of the Jahn–Teller effect. The primary confirming features in the spin resonance are summarized as follows:

(1) Analysis of the resolved Si^{29} hyperfine interactions indicates that $\sim$60% of the wave function is localized in a dangling bond on a single silicon atom. An additional 20–25% of the wave function can be accounted for by partially resolved hyperfine interactions on a few (5-6) near-neighbor sites, with the remaining 15–20% presumably spread out over more distant sites.

(2) The spectrum also reveals resolved hyperfine interaction with a 100% abundant $I=\frac{1}{2}$ impurity. ENDOR measurements confirm that this impurity is phosphorus. Only $\sim$1% of the wave function is localized on this atom.

(3) The g tensor is nearly axially symmetric reflecting the axis of the dangling bond which makes up most of the wave function. Its anisotropy is similar to other defects which we have studied and have identified as made up of dangling bonds associated with vacancies. Also, a simple "localized molecule" treatment of the defect has been given which indicates that the magnitude and sign of the g shifts are reasonable.

(4) Motional broadening and narrowing effects are observed in the spectrum in the temperature region 60–150°K and are attributed to the thermally activated reorientation from one Jahn-Teller distortion direction to the other. Analysis confirms that the motion is that of an electronic "bond switching" (not atom motion) and that the motion is around the phosphorus-vacancy axis as predicted by the model.

(5) This motion is also studied by stress induced alignment at 20.4°K. Combined with the linewidth studies, the kinetics of the reorientation process can be obtained with very high precision. The observed activation energy $\sim$0.06 eV appears reasonable for this process. The *sense* of the alignment is consistent with the natural preference of the defect to align so that the applied stress aids the Jahn-Teller distortion. The *degree* of alignment can be accounted for by a simple model which only considers the change in energy per unit strain along the axis joining the bonding Si-Si pair. The magnitude deduced for this parameter is reasonable, being of the order of observed deformation potentials.

(6) Stress-induced alignment of the phosphorus-vacancy axis is also studied. The sense and magnitude can also be accounted for in terms of the model. The kinetics of this process, with an activation energy of 0.93 ± 0.05 eV, appear reasonable for the atom rearrangements required by the model.

The E center, in addition to being an important defect in radiation damage studies, may play an important role in unirradiated silicon. High-temperature diffusion of substitutional impurity atoms, as well as self-diffusion, is generally believed to be via lattice vacancies. As a result, such impurity-vacancy interactions are important in understanding the mechanics of impurity diffusion. The E center represents a unique opportunity to probe this interaction in detail. Being the neutral charge state appropriate for the Fermi level in the middle of the gap, it is in the correct charge state relevant to the high-temperature experiments. In Fig. 18, we summarize what we have been able to learn about this interaction. The abscissa denotes schematically the different neighboring atom positions as one moves out from the phosphorus. They are not necessarily in a straight line. We don't know the exact shape to be expected but it seems reasonable to assume it is something as shown. From the kinetics of the stress-orientation of the phosphorus-vacancy axis we have determined the height of the second maximum to be 0.93 ± 0.05 eV, as indicated. We have also made two estimates for the indicated difference in barrier heights ΔE from annealing studies. The smaller estimate comes from EPR measurements and the larger one comes from electrical measurements. From previous studies[36] we estimate that the activation energy for the isolated vacancy motion is[41] $E_M{}^V=0.33\pm0.20$ eV. The indicated estimate of the binding energy $E_C\sim0.28$ eV at the third-neighbor site comes from a simple calculation of the electrostatic interaction between a negative vacancy at this site and a positive phosphorus atom. This value is seen to be generally consistent with the other experimentally determined estimates.

We have therefore been able to deduce quite a bit about the details of the phosphorus-vacancy interaction. It is hoped that future experiments can narrow further

[41] The value for p-type silicon has been determined to be (Ref. 36) 0.33 ± 0.03 eV. However, in annealing studies in n-type silicon (Refs. 3, 36), it is found that both A and E centers emerge in two discrete temperature regions. One of these is presumably to be associated with long-range motion of the vacancy, and, depending upon which we select, we either deduce $\sim0.1-0.2$ eV or $\sim0.4-0.5$ eV for the activation energy of the vacancy motion. This difference from the p-type value presumably reflects the role of charge state in the mobility of the vacancy. Since we do not know the charge state appropriate to Fig. 18, we indicate an error ±0.2 eV to bracket these extremes.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the errors in $E_M{}^V$ and ΔE with the result that this interaction will eventually be rather precisely mapped out. However, even with the present uncertainties we can draw certain conclusions:

From Fig. 18, we see that there is a sizeable binding energy ($\sim$1.0 eV) between the phosphorus atom and a vacancy in their closest bound state.[42] Because of this attractive interaction there is an enhanced probability of finding a vacancy near the phosphorus at elevated temperatures and we therefore expect the diffusion of the phosphorus to be enhanced over that for silicon self-diffusion. However, we recognize that *it is only the binding at the third-atom site that is operative in enhancing the diffusion* because the vacancy must go through this site in order for the phosphorus atom to diffuse.[43] For this binding energy, we take the Coulomb estimate E_C=0.28 eV. We estimate the error in this quantity somewhat arbitrarily as $\pm$0.15 eV, which we feel, in view of the general consistency of the values in Fig. 18, is overly cautious.

Phosphorus diffusion in silicon has been measured and the activation energy has been determined to be[44] 3.66$\pm$0.18 eV. However, there has been no corresponding measurement of self-diffusion in silicon. Using the value 0.28$\pm$0.15 eV for the lowering of the phosphorus activation energy, we therefore estimate the activation energy for silicon self-diffusion

$$E_{\text{self-diff.}}=3.94\pm0.33 \text{ eV}. \quad (27)$$

Since this quantity has not been measured directly, we feel that this probably represents the best present estimate that can be made for this important quantity.

Combining this with the activation energy for vacancy motion of 0.33$\pm$0.20 eV, we obtain for the formation energy of the lattice vacancy[45]

$$W=3.6\pm0.5 \text{ eV}. \quad (28)$$

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the assistance of W. Colliton in all phases of the measurements. Helpful conversations with Dr. F. Ham, Dr. B. Segall, Dr. R. W. Redington, and Dr. R. M. Walker are also gratefully acknowledged.

APPENDIX A: ESTIMATE OF $|\psi_{3s}(0)|^2$ AND $\langle r_{3p}{}^{-3}\rangle$ FOR THE NEUTRAL $3p$ ATOMS

In previous publications,[1,4] we have used $|\psi_{3s}(0)|^2_{Si}\approx24(10^{24})$ cm^{-3} and $\langle r_{3p}{}^{-3}\rangle_{Si}\approx17(10^{24})$ cm^{-3} in the analysis of resonance centers. These estimates were originally made using various Hartree functions, available only for higher ionization states of silicon, as guides. Since then, Hartree-Fock calculations for neutral silicon, as well as all of the other neutral $3p$ atoms, have been reported by Watson and Freeman[15] and a revised estimate of these quantities is in order.

In Table III, the first two columns show $|\psi_{3s}(0)|^2$ and $\langle r_{3p}{}^{-3}\rangle$ determined from the Watson and Freeman wave functions[46] for the atoms aluminum through chlorine. Consider first $|\psi_{3s}(0)|^2$. As was pointed out by Kohn and Luttinger,[47] direct comparison with experiment in the case of the alkalies shows that Hartree-Fock wave functions generally underestimate $|\psi_s(0)|^2$, and the true values will probably be somewhat larger than these values. In the case of silicon, Kohn and Luttinger estimated an appropriate scaling factor by comparing the Hartree-Fock estimate of $|\psi_{3s}(0)|^2$ for the alkali-like Si^{3+} ion[48] with that predicted by the Goudsmit formula[49] for such ions and assuming the same scaling factor to apply for Si^0. Their result was that the Hartree-Fock value should be increased by $\sim$19%. Applying the same correction to the Watson and Freeman value for Si^0, we obtain the final estimate of 31.5(10^{24}) cm^{-3} shown in the table.

Now consider $\langle r_{3p}{}^{-3}\rangle$. Barnes and Smith[50] have estimated this quantity from an analysis of optical fine structure splittings and their values are given in column three of Table III for each of the $3p$ atoms. In the case of aluminum[51] and chlorine,[52] atomic beam hyperfine measurements give accurate determinations of $\langle r_{3p}{}^{-3}\rangle$, also shown in the table, to which we can compare the estimates of Barnes and Smith. We see that the agreement is very good, the atomic beam values lying $\sim$3% higher. We thus increase the Barnes and Smith estimates by 3% for the remaining atoms to give the final estimate of $\langle r_{3p}{}^{-3}\rangle$ shown in the last column.

We note that the $\langle r_{3p}{}^{-3}\rangle$ determined from the Watson and Freeman wave functions are uniformly too low. It is interesting to note that for Si^0, the value must be increased by 19% to agree with our final estimate. Although this agreement with the scaling factor for

[42] We remind the reader that these arguments (and the values in Fig. 18) apply only to the neutral charge state of the defect and therefore when the Fermi level is near the middle of the gap. This should apply to the high temperatures used in diffusion studies for anything but the most heavily doped materials.

[43] This important fact (see Sec. IVC and Fig. 17) has been generally overlooked by previous authors.

[44] C. S. Fuller and J. A. Ditzenberger, J. Appl. Phys. 27, 544 (1956).

[45] We have, of course, assumed throughout this discussion that both phosphorus and self-diffusion proceed by the vacancy mechanism. If other mechanisms are also involved, the formation energy will be even larger.

[46] We are indebted to B. Segall for supplying us with tabulated values of the $3p$ Watson and Freeman wave functions from which $\langle r_{3p}{}^{-3}\rangle$ were calculated.

[47] W. Kohn and J. M. Luttinger, Phys. Rev. 97, 883 (1961).

[48] W. Hartree, D. R. Hartree, and M. F. Manning, Phys. Rev. 60, 857 (1941).

[49] S. Goudsmit, Phys. Rev. 43, 636 (1933).

[50] R. G. Barnes and W. V. Smith, Phys. Rev. 93, 95 (1954).

[51] H. Lew and G. Wessel, Phys. Rev. 90, 1 (1953).

[52] V. Jaccarino and J. G. King, Phys. Rev. 83, 471 (1951); J. G. King and V. Jaccarino, *ibid.* 84, 852 (1951). $\langle r_{3p}{}^{-3}\rangle$ was determined from their results by analysis similar to that in the preceding Ref. 51.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Various estimates of $|\psi_{3s}(0)|^2$ and $\langle r_{3p}^{-3}\rangle$ for the neutral $3p$ atoms. The indicated values were determined from (HF) Hartree-Fock wave functions, (FS) experimental fine structure constants, and (AB) atomic-beam hyperfine measurements. See text for the method of determining the final estimates.

Atom	$\lvert\psi_{3s}(0)\rvert^2_{\rm HF}$ (10^{24} cm^{-3})	$\langle r_{3p}^{-3}\rangle_{\rm HF}$ (10^{24} cm^{-3})	$\langle r_{3p}^{-3}\rangle_{\rm FS}$ (10^{24} cm^{-3})	$\langle r_{3p}^{-3}\rangle_{\rm AB}$ (10^{24} cm^{-3})	$\lvert\psi_{3s}(0)\rvert^2$ Scaling factor	Final Estimates $\lvert\psi_{3s}(0)\rvert^2$ (10^{24} cm^{-3})	Final Estimates $\langle r_{3p}^{-3}\rangle$ (10^{24} cm^{-3})
Al	15.8[a]	6.94[a]	8.64[b]	8.95[c]	1.29	20.4	8.95
Si	26.4[a]	13.5[a]	15.6[b]	...	1.19	31.5	16.1
P	37.8[a]	22.0[a]	23.5[b]	...	1.10	41.6	24.2
S	53.3[a]	31.9[a]	33.9[b]	...	1.09	58.4	34.9
Cl	71.8[a]	44.2[a]	48.3[b]	49.5[d]	1.12	80.4	49.5

[a] Calculated from wavefunctions in R. E. Watson and A. J. Freeman, Phys. Rev. 123, 521 (1961).
[b] R. G. Barnes and W. V. Smith, Phys. Rev. 93, 95 (1954).
[c] H. Lew and G. Wessel, Phys. Rev. 90, 1 (1953).
[d] V. Jaccarino and J. G. King, Phys. Rev. 83, 471 (1951); J. G. King and V. Jaccarino, *ibid.* 84, 852 (1951).

$|\psi_{3s}(0)|^2$ must be considered somewhat fortuitous, it does suggest that both $\langle r_{3p}^{-3}\rangle$ and $|\psi_{3s}(0)|^2$ are underestimated by a comparable amount in these Hartree-Fock wave functions. This then suggests a consistent way to determine the proper $|\psi_{3s}(0)|^2$ scaling factor for the other atoms: For the atoms other than Si, the scaling factor was determined as being equal to the ratio of our final estimate of $\langle r_{3p}^{-3}\rangle$ to that determined from the Freeman-Watson wave functions. The scaling factor thus determined is shown in Table III along with the corresponding final estimate of $|\psi_{3s}(0)|^2$ for all the $3p$ neutral atoms.[53]

We will not attempt to set limits of accuracy on the values of Table III. There are still sufficient uncertainties in the applicability of these free atom values to those of a bonded atom surrounding a defect in a solid to make such speculations somewhat academic for our use, anyway. However, Table III does give a set of reasonable values, consistently arrived at, for the complete $3p$ series of neutral atoms.

In this paper we have used only the values for silicon and phosphorus. In subsequent papers in this series, we will have occasion to use the estimates for others of the atoms in the table.

APPENDIX B: APPROXIMATE LCAO TREATMENT OF THE g SHIFT

We approximate the one-electron LCAO molecular orbitals of Fig. 8 as:

$$\bar{\psi}'(A)=\sigma_3{}^{I}(A)=\alpha_I\psi_{3s}(A)+\beta_I\psi_{3pI}(A),$$

$$\bar{\psi}^{\pm}(AX_2)=N^{\pm}[\sigma_3{}^{II}(A)\pm\sigma_3{}^{-II}(X_2) \mp\sum_i \epsilon_i\{\psi_i(A)\pm\psi_i(X_2)\}], \quad \text{(B1)}$$

with similar expressions for the other orbitals. Here $\sigma_3{}^{\rm I}$, $\sigma_3{}^{\rm II}$, $\sigma_3{}^{\rm III}$, and $\sigma_3{}^{\rm IV}$ represent four orthogonal cylindrical orbitals made up of $3s$-$3p$ atomic wave functions, where the superscript represents the axis of the directed orbital as shown in Fig. 8. (By $\sigma_3{}^{-\rm II}$, we denote an orbital pointed in the opposite direction from $\sigma_3{}^{\rm II}$.) We have included small admixtures of the core wave functions ψ_i, ($1s$, $2s$, and $2p$), in order to orthogonalize the total wave function to the core. The small admixture coefficients are given by

$$\epsilon_i=\langle\sigma_3{}^{-\rm II}(X_2)|\psi_i(A)\rangle, \quad \text{etc.}$$

The normalization coefficients are

$$N^{\pm}=[2(1\pm S-\sum_i \epsilon_i{}^2)]^{-1/2},$$

with

$$S=\langle\sigma_3{}^{\rm II}(A)|\sigma_3{}^{-\rm II}(X_2)\rangle.$$

A typical orbital angular momentum matrix element of (15) is

$$\langle\bar{\psi}^{\pm}(AX_2)|L_j|\bar{\psi}'(A)\rangle = \sum_k \langle\bar{\psi}^{\pm}(AX_2)|\sigma_3{}^{k}(A)\rangle\langle\sigma_3{}^{k}(A)|L_j|\sigma_3{}^{I}(A)\rangle.$$

Substituting the expression for $\bar{\psi}^{\pm}(AX_2)$ given by Eq. (B1), and using the fact that only $\sigma_3{}^{\rm II}(A)$ has significant overlap with any of the terms for $\bar{\psi}^{\pm}(AX_2)$, we obtain

$$\langle\bar{\psi}^{\pm}(AX)_2|L_j|\bar{\psi}'(A)\cong\tfrac{1}{2}(N^{\pm})^{-1}\langle\sigma^{\rm II}(A)|L_j|\sigma^{\rm I}(A)\rangle. \quad \text{(B2)}$$

Now consider the corresponding matrix element in (15) involving $\mathbf{V}_{so}$

$$\langle\bar{\psi}'(A)|(\mathbf{V}_{so})_i|\bar{\psi}^{\pm}(AX_2)\rangle = (\beta/mc)\langle\sigma_3{}^{\rm I}(A)|(\mathbf{E}\times\mathbf{p})_i|\bar{\psi}^{\pm}(AX_2)\rangle. \quad \text{(B3)}$$

The electric field $\mathbf{E}$ is large only near the nuclei of atoms A and X_2 and therefore, in these matrix elements, tends to weight strongly only the parts of the wave functions near the atomic cores. To a good approximation we may therefore retain in Eq. (B1) only the atomic terms centered on atom A giving

$$\cong(\beta/mc)N^{\pm}\langle\sigma_3{}^{\rm I}(A)|(\mathbf{E}\times\mathbf{p})_i|\sigma_3{}^{\rm II}(A)-\epsilon_{2p}\psi_{2p\rm II}(A)\rangle.$$

[53] The scaling factor is not constant throughout the series but is largest at the beginning. A further check that this variation is not unreasonable is given by noting that an even larger scaling factor of 1.42 is required in the case of neutral sodium (which starts this row in the periodic table), in order to raise the value determined from the Hartree-Fock calculations [D. R. Hartree and W. Hartree, Proc. Roy Soc. (London) A193, 299 (1948)] (3.56×10^{24} cm^{-3}) to that observed experimentally [P. Kusch and H. Taub, Phys. Rev. 75, 1477 (1949)] (5.05×10^{24} cm^{-3}).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(The terms centered on atom X_2, though not zero near atom A, are slowly varying there, and therefore do not contribute significantly to the spin-orbit interaction.) Having reduced the matrix elements to those on atom A only, we may now make the substitution $\mathbf{E}=[\mathbf{r}/r]E(r)$ for the central force field near the nucleus of atom A, leading to the familiar free atom expressions

$$\langle nlm|(\mathrm{V}_{so})_i|n'l'm'\rangle=\delta_{ll'}\lambda_{nl,n'l}\langle lm|L_i|lm'\rangle,$$

with

$$\lambda_{nl,n'l}=(\beta/mc)\langle nl|[E(r)/r]|n'l\rangle.$$

This gives, upon straightforward manipulation,

$$\langle\bar{\psi}'(A)|(\mathrm{V}_{so})_i|\bar{\psi}^{\pm}(AX_2)\rangle \cong N^{\pm}[\lambda_{3p}\mp(\epsilon_{2p}\lambda_{2p,3p})/\beta_I]\langle\sigma^{\mathrm{I}}(A)|\mathbf{L}_i|\sigma^{\mathrm{II}}(A)\rangle. \quad \text{(B3')}$$

Substituting (B2) and (B3') into (15), and summing over-all excited orbitals in (B1),

$$\Delta g_{\mathrm{II}}=0,$$

$$\Delta g_{\perp}\cong\lambda_{3p}\beta_I{}^2\left[\frac{1+\gamma}{E_b}\quad\frac{1-\gamma}{E_a}\right], \quad \text{(B4)}$$

with

$$\gamma=-\frac{\lambda_{2p,3p}\epsilon_{2p}}{\lambda_{3p}\beta_I}. \quad \text{(B5)}$$

Here, γ is a positive number because $\lambda_{2p,3p}$ and ϵ_{2p} will always be of opposite sign, a consequence of the fact that ψ_{3p} has one more radial node than ψ_{2p}. [In our previous treatment (I), the core overlap terms were not included in the wave functions. For these wave functions, $\gamma=0$, and (B4) leads to Eq. (16) in the text of the present paper.]

We may estimate γ, as follows: We approximate $\lambda_{2p,3p}$ as $-(\lambda_{2p}\lambda_{3p})^{1/2}$ using the values $\lambda_{2p}=0.40$ eV estimated by Liu[54] from x-ray emission data[55] and $\lambda_{3p}=0.02$ eV.[23] We calculate ϵ_{2p} from Eq. (19) using Slater atomic orbitals[56] and standard overlap integral formulas[57] in conjunction with certain required integrals tabulated by Rosen.[58] For the normal Si-Si lattice distance and with the $\sigma^i(A)$ taken as the tetrahedral sp^3 hybrids (25% $3s$, 75% $3p$), we obtain $\epsilon_{2p}=0.033$. With $\beta_I=(3)^{1/2}/2$ for the tetrahedral hybrid, this gives $\gamma\cong 0.17$.

---

[54] L. Liu, Phys. Rev. 126, 1317 (1962).
[55] D. H. Tombulian and W. M. Cadis, Phys. Rev. 59, 422 (1944).
[56] J. C. Slater, Phys. Rev. 36, 57 (1930).
[57] R. S. Mulliken, C. A. Rieke, D. Orloff, and H. Orloff, J. Chem. Phys. 17, 1248 (1949).
[58] N. Rosen, Phys. Rev. 38, 255 (1931).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Experimental Investigations of Single Injection in Compensated Silicon at Low Temperatures*

Bob L. Gregory† and Angel G. Jordan
Electrical Engineering Department, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received 23 December 1963; revised manuscript received 24 February 1964)

Experimental investigations of the phenomenon of single carrier, space-charge limited injection in silicon at liquid-helium temperatures have been performed in an attempt to validate the model for space-charge limited currents (SCLC) in an insulator with traps. Comparisons between predicted and experimental characteristics were performed on p_+pp_+ devices fabricated by diffusion techniques. Pulse and dc measurements were made to obtain steady-state current-voltage characteristics for the above devices. Modifications were observed in the SCLC characteristics due to field-dependent carrier mobilities, the exact nature of which could be determined from the experimental steady-state characteristics. The density of deep trapping centers in the above material at liquid-helium temperature, corresponding to the compensation, was obtained by measurements of the traps-filled voltage. An independent value for the compensation in the same material was obtained by Hall measurements in the temperature range $20°K<T<77°K$. Transient measurements of the single-carrier current were performed, from which values for the capture probability and capture cross section for the deep trapping center could be obtained. A comparison was made between the capture cross section obtained in this manner and the theoretical prediction of the "giant trap" model.

INTRODUCTION

THE theory of space-charge limited currents (SCLC) in insulators with traps as proposed by Lampert[1] has been employed frequently in recent years to explain the current-voltage characteristics of insulator structures. In some instances, however, uncertainty arose in the interpretation of the insulator characteristics because of apparent disagreement between material parameters obtained from the single injection interpretation and independent material measurements. For example, in CdS, single-carrier SCLC was investigated by Bube[2] in an attempt to correlate the trap density with the traps filled voltage (V_{TFL}) as defined by Lampert. However, his determinations of trap density obtained from measurements of the apparent traps filled voltage differed by as much as three orders of magnitude from independently obtained trap-density values.

This experimental uncertainty in the model for SCLC in insulators with traps prompted the work in this paper. It was observed, originally by Lampert, that the purest insulator in which to test the model would be a high-purity semiconductor at low temperatures (sufficiently low to cause the equilibrium carrier density to be completely negligible). Since the dopant densities in the semiconductor would be known (or could be determined by Hall effect measurements), one should be able to quantitatively compare the predictions of the theoretical model with the experimental characteristics. Other important results expected from investigation of single injection at low temperatures are:

(1) the capture probability, hence capture cross section, of the deep trapping levels present due to compensation could be measured directly by pulse techniques;

(2) the field dependence of the drift mobility could be determined from the SCLC characteristic in the Child's law region;

(3) the injecting nature of contacts at low temperatures could be explored; and

(4) the nature of scattering mechanisms at low temperatures could be better understood by interpreting the field-dependent mobility data.

p_+pp_+ SILICON DEVICES AT LOW TEMPERATURES

As mentioned above, the ideal material in which to examine the model for single-carrier SCLC should be a high-purity semiconductor at low temperatures. In actuality, however, one is limited to germanium or silicon since no other sufficiently pure semiconductors are available. Furthermore, high-purity germanium is unsuitable for these investigations since at low temperatures impact ionization occurs at quite low field strengths,[3,4] destroying its insulating nature. Thus, presently only silicon is suitable for investigating SCLC at low temperature; furthermore, since highest purity silicon is invariably p type due to the presence of boron, p-type material was chosen for the experiment in this work.

As the temperature of a sample is reduced to below the extrinsic range ($T<50°K$ in lightly doped material) the free holes, produced by thermal ionization of noncompensated acceptor centers, recombine with ionized acceptors producing neutral centers. At 4.2°K (liquid-helium temperature) virtually all free holes will have disappeared and N_A-N_D neutral acceptor centers will have appeared. The electrical effect of the compensated

* This work was supported in part by the National Science Foundation.
† Present address: Sandia Corporation, Sandia Base, Albuquerque, New Mexico.
[1] M. A. Lampert, Phys. Rev. 103, 1648 (1956); J. Appl. Phys. 29, 1082 (1958).
[2] R. H. Bube, J. Appl. Phys. 33, 1733 (1962).
[3] N. Sclar and E. Burstein, Phys. Rev. 98, 1757 (1955).
[4] S. H. Koenig, R. D. Brown, and W. Schillinger, Phys. Rev. 128, 1668 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

centers at low temperature is to produce trapping levels for injected carriers. The positively charged donor centers would be trapping levels for injected electrons. Similarly, the negatively charged acceptor centers would be traps for injected holes. In this simplified model the impurities consist of N_A shallow acceptors and N_D shallow donors (shallow at room temperature) where $N_A > N_D$. The trapping kinetics for this model depends in large part on the "capture cross section" of the ionized centers. The theoretical considerations which allow this cross section to be determined will be discussed in a later section dealing with the cascade capture process.[5]

To fabricate a single-carrier SCLC device, one must determine which actual contacts exhibit the properties assumed in the model. To reiterate these properties, one contact must be "ohmic" and the other "blocking." At the ohmic contact a large reservoir of holes which can be readily injected into the bulk under forward bias must exist. Such contact is not easily achieved for temperatures in the liquid-helium range. For example, it is known that most metal contacts to silicon create Schottky-type barriers at the metal-semiconductor interface which, though small at room temperature in some cases, are extremely large compared to kT at 4.2°K. Such a barrier would, of course, prevent hole injection at this temperature. A nondegenerate diffused or alloyed p contact would also be unsuitable for use as an injecting contact at these low temperatures since the contact region would merely be an extension of the insulating bulk. However, a degenerate p_+ contact appears to be suitable for the injecting contact at low temperatures since in the degenerate p_+ region there exists a large number of mobile holes which could easily be injected into the insulating region because only a slight barrier exists.

The exact nature of the blocking contact at low temperatures is not so important. Any contact which prevents neutralizing electrons from entering when holes are being injected at the ohmic contact is satisfactory. For this investigation of SCLC the most suitable contact would again be a degenerate p_+ contact since no rectification is desired. No electrons can be injected at this contact at low temperatures regardless of the bias polarity.

The structures in which SCLC was explored at low temperatures were, therefore, p_+pp_+ devices with degenerate p_+ regions. The schematic potential solution for the entire structure in equilibrium is shown in Fig. 1. The metallurgical junctions occur at $X=0$, d in this diagram.

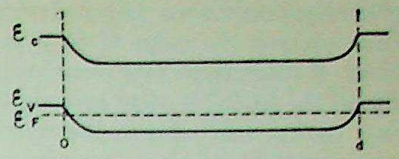


FIG. 1. p_+pp_+ structure in equilibrium.

FIELD-DEPENDENT MOBILITIES AT LOW TEMPERATURES

At temperatures in the liquid-helium range carrier mobilities become field-dependent for values of electric field orders of magnitude smaller than at room temperature. To obtain an explanation for this phenomenon, one must consider the temperature dependence of the mobility. Since at 4.2°K the low-field mobility in lightly doped silicon is more than two orders of magnitude greater than its room-temperature value,[6] the carrier drift velocity at this temperature for a given field strength is likely to greatly exceed the room temperature drift velocity at the same field strength. Therefore the breakdown of the constant mobility assumption, which occurs when the energy given to mobile carriers by the field between collisions is comparable to the equilibrium carrier energy, should occur at greatly reduced electric fields at low temperatures due both to the increased mobility and to the decreased equilibrium carrier energy at these temperatures.

The two principal mechanisms which determine the form of the mobility field dependence are acoustical and optical scattering. Only at very low fields must ionized impurity scattering be considered.[7]

The equation governing the dependence of the mobility μ and electric field E when only acoustical scattering is of importance is

$$\mu = (8/3)^{1/4}(\mu_{LOC})^{1/2}E^{-1/2}, \qquad (1)$$

where μ_{LO} is the acoustical scattering limited zero-field mobility and c is the velocity of sound in the semiconductor.

The field strength at which optical scattering should become important in determining the mobility can be approximated by employing the above equation. If one equates the carrier drift velocity as calculated from this equation to the saturated drift velocity as obtained by equating the carrier kinetic energy to the Raman energy $\mathcal{E}_R$ the following results,

$$(8/3)^{1/4}(\mu_{LOC})^{1/2}E^{1/2} \cong 1/2(2\mathcal{E}_R/m^*)^{1/2}, \qquad (2)$$

where m^* is the density-of-states effective mass.

When the appropriate constants are inserted, it is obtained that in p-type silicon for $E > 50$ V/cm optical scattering should determine the mobility at 4.2°K. This result is surprising when one considers that electric fields near 10^5 V/cm were required at room temperature[8] to cause the carrier drift velocity to approach saturation.

As shall be seen in the following section, the minimum electric field value for which SCLC is observed in the

[5] M. Lax, Phys. Rev. 119, 1502 (1960).

[6] E. H. Putley and W. H. Mitchell, Proc. Phys. Soc. (London) 72, 193 (1958).

[7] J. B. Gunn, *Progress in Semiconductors* (John Wiley & Sons, Inc., New York, 1957), Vol. 2, pp. 213–247.

[8] A. C. Prior, Phys. Chem. Solids 12, 175 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

present compensated silicon devices at low temperatures exceeds the value of 50 V/cm by more than an order of magnitude. One would expect, therefore, that for all the SCLC devices investigated at low-temperature optical scattering is the process determining the carrier mobility.

The simple model for optical scattering, which assumes all carriers are scattered optically upon reaching the threshold energy for optical-phonon emission, allows no carrier energies higher than the threshold energy and predicts drift velocity saturation. When higher carrier energies are allowed, as would be in any realistic model, the drift velocity would not actually saturate. If, for example, the mobility varied as $E^{-\alpha}$, where $\alpha<1$, as some workers have reported, the drift velocity would be proportional to $E^{1-\alpha}$. If one assumes that due to optical scattering the hole mobility μ_P is given by the expression

$$\mu_P=K_0E^{-\alpha}, \tag{3}$$

the constants K_0 and α can be evaluated empirically by comparing the observed Child's law region characteristics to the theoretical Child's law characteristic. When the mobility variation of Eq. (3) is assumed, one can calculate the Child's law characteristic. The resulting equation relating the current density j to the terminal voltage V is

$$j=[(3-\alpha)/(2-\alpha)]^{2-\alpha}(\epsilon K_0)/(2-\alpha)V^{(2-\alpha)}/d^{(3-\alpha)}, \tag{4}$$

where ϵ is the permittivity of the material and d is the device thickness. This equation can be seen to be correct by allowing α to equal zero. In this case the equation reduces to the familiar square-law characteristic for constant mobility.

TRAPPING AT LOW TEMPERATURES

At low temperatures in p_+pp_+ single injection devices, the hole trapping process should involve "giant" traps[5] since Coulombic attraction exists between injected holes and the negatively charged compensated boron atoms. An injected hole would therefore be trapped first in a large orbit excited state of a negative boron center via the emission of a single acoustical or optical phonon. After being trapped in this excited state, the hole would decay to the ground state of the center via a cascade of single-phonon processes.

Whenever carriers are injected into insulators where deep trapping centers exist, the rate at which the carriers are trapped is naturally related to the capture cross section of the centers. In the case of holes injected into p type silicon, the rate at which the carriers are trapped is given by the equation

$$dp/dt=-\bar{B}pN_A^-, \tag{5}$$

where N_A^- is the density of ionized acceptors, p is the density of injected holes, and $\bar{B}$ is the mean-capture probability. $\bar{B}$ is obtained by averaging $B(\mathcal{E})$ over the energy distribution of the injected carriers. The energy-dependent capture probability can be written as

$$B(\mathcal{E})=\sigma(\mathcal{E})(2\mathcal{E}/m^*)^{1/2}, \tag{6}$$

where $\sigma(\mathcal{E})$ is the energy-dependent cross section.

The mean capture probability is therefore seen to be

$$\bar{B}=\langle\sigma(\mathcal{E})(2\mathcal{E}/m^*)^{1/2}\rangle. \tag{7}$$

The experimental values of $\bar{B}$, from which estimates for the capture cross section can be made, can be obtained by comparing the experimental current transients with the analytically derived characteristic. For the condition of slow trapping (the drift transit time $t_d\ll\tau_t$, the trapping time) the following analysis may be employed to obtain the analytical characteristic.

TRANSIENT SCLC IN THE PRESENCE OF SLOW TRAPPING

Carrier flow in a plane parallel insulator with deep traps can be described with four equations: the transport equation, the continuity equation, Poisson's equation, and a rate equation describing the trapping process. If diffusion currents are again neglected, these four equations are as written below:

$$j(X,t)=q\mu_pp(X,t)E(X,t); \tag{8}$$

$$\frac{\partial j(X,t)}{\partial X}=-q\left[\frac{\partial p(X,t)}{\partial t}+\frac{\partial p_t(X,t)}{\partial t}\right]; \tag{9}$$

$$\frac{\partial E(X,t)}{\partial X}=\frac{q}{\epsilon}[p(X,t)+p_t(X,t)+N_D-N_A-n-n_t]; \tag{10}$$

and

$$\frac{\partial p_t(X,t)}{\partial t}=\bar{B}p(X,t)[N_T-p_t(X,t)], \tag{11}$$

where q is the electronic charge, p_t the density of trapped holes, n the density of mobile electrons, n_t the density of trapped electrons, and N_T the density of trapping centers.

In writing Eq. (11), the generation of carriers from the deep traps has been assumed to be negligible. Additional assumptions which appear to be appropriate to the low-temperature situation are:

(1) due to the quasistationary[9] behavior of the carrier flow in the slow trapping situation, both (a) the current divergence $\partial j(X,t)/\partial X$ and (b) displacement currents may be neglected;

(2) at 4.2°K, both n and n_t are negligibly small in p-type material;

(3) since the injected holes are captured on acceptor centers, N_T is equal to N_A in Eq. (11); and

(4) since $t_d\ll\tau_t$, the trap free SCLC solution is assumed to apply at $t=0$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The simplified equations can be rewritten as

$$j(t)=q\mu_p p(X,t)E(X,t), \tag{12}$$

$$\frac{\partial p(X,t)}{\partial t}=-\frac{\partial p_t(X,t)}{\partial t}, \tag{13}$$

$$\frac{\partial E(X)}{\partial X}=\frac{q}{\epsilon}[p(X,t)+p_t(X,t)+N_D-N_A], \tag{14}$$

$$\frac{\partial p_t(X,t)}{\partial t}=\bar{B}p(X,t)[N_A-p_t(X,t)]. \tag{15}$$

To solve this system of equations for $j(t)$, one proceeds as follows:

(1) relate $p_t(X,t)$ and $p(X,t)$ via Eqs. (13) and (15);

(2) relate $j(t)$ to $p_t(X,t)$ using these results and Eq. (12); and

(3) substitute SCLC initial conditions into the problem obtaining the solution for $j(t)$.

The resulting expression for $j(t)$ is

$$j(t)=J_0\left[\frac{1-[2(2-\alpha)^2(d/X)^{(\alpha-1)/(2-\alpha)}V_{\mathrm{TFL}}/(3-\alpha)V]}{1-[2(2-\alpha)^2(d/X)^{(\alpha-1)/(2-\alpha)}V_{\mathrm{TFL}}/(3-\alpha)V]\exp\bar{B}tN_D\{1-(3-\alpha)(X/d)^{(\alpha-1)/(2-\alpha)}V/2(2-\alpha)^2V_{\mathrm{TFL}}\}}\right]. \tag{16}$$

Equation (16) does not agree completely with the assumptions made in its derivation since it predicts a nondivergenceless current. The divergence is not great, however, since the exponent of the X terms, $(\alpha-1)/(2-\alpha)$, is much less than one for values of α appropriate to the devices of the following section. In fact, for $\alpha=1$, corresponding to drift-velocity saturation, the X dependence disappears altogether yielding a completely divergenceless current. Since the assumptions upon which the present analysis is based are strictly valid only for the case of drift-velocity saturation, the theoretical current transients shall be calculated under this assumption. The difference in form between the characteristics obtained in this manner and the results of an exact solution should not be great since, for values of α reasonably close to one, the X dependence of the current in Eq. (16) is greatly depressed. Hence, by matching the theoretical current transients obtained assuming drift-velocity saturation to the experimental characteristics, one should be able to obtain approximate values for $\bar{B}$ even though the drift velocity has not reached saturation. The theoretical transient characteristic to be employed in these comparisons is obtained by substituting $\alpha=1$ into Eq. (16). This yields

$$j(t)=J_0\left[\frac{1-V_{\mathrm{TFL}}/V}{1-(V_{\mathrm{TFL}}/V)\exp\{\bar{B}tN_C(1-V_{\mathrm{TFL}}/V)\}}\right]. \tag{17}$$

Several limit checks may be performed on this solution to compare it to reality. For example, an obvious limit check is an examination of the behavior of the equation as $t\rightarrow\infty$ for the two cases of $V<V_{\mathrm{TFL}}$ and $V>V_{\mathrm{TFL}}$. For $V<V_{\mathrm{TFL}}$, Eq. (17) predicts that $j(t)\rightarrow 0$ as $t\rightarrow\infty$. This corresponds to the result expected from physical considerations. For the second situation where $V>V_{\mathrm{TFL}}$, the equation predicts that $j(t)\rightarrow J_0(1-V_{\mathrm{TFL}}/V)$, which also corresponds to the expected behavior. An additional check on the equation may be obtained by allowing the density of compensation N_D to become very large. In this case, V_{TFL}/V would be much greater than 1 and $j(t)\rightarrow J_0\exp(-\bar{B}tN_d)$. This simple exponential is the transient characteristic one would expect from physical reasoning for this case, since the density of ionized trapping levels is so large as to remain essentially constant even though carriers are being trapped.[9]

DEVICE CONSTRUCTION AND EVALUATION

The degenerate p_+ contacts were formed by comparatively low-temperature boron diffusions. For the temperatures employed (900–950°K) the surface concentration of boron exceeded 10^{20} cm^{-3} as was observed from sheet resistivity and junction depth measurements on identically diffused n-type wafers.[10] These measurements indicated junction depths of approximately 1 μ in addition to the above value of surface concentration.

Two techniques were employed to fabricate single injection devices from the diffused wafers. Circular devices with areas of approximately 10^{-2} cm^2 were obtained by masking and etching. Rectangular devices of comparable area were made by scribing and breaking the diffused samples. Leads were attached to the above devices by thermocompression bonding aluminum wire.

To obtain measurements at 4.2°K on these devices, a shielded, light-tight cryostat was employed. To obtain the necessary low temperature, the cryostat was inserted into the liquid-helium chamber of a double-walled transport Dewar.

To measure the steady-state current-voltage characteristics over the relatively large range of currents observed in SCLC devices, several techniques were employed. Below 1 mA, dc measurements could be taken without appreciable heating. Consequently currents between 10^{-9} and 10^{-3} A were observed using dc schemes. Point by point measurements were made in some cases using a well-shielded circuit employing a

[9] A. Many and G. Rakavy, Phys. Rev. 126, 1980 (1962).
[10] G. Backenstoss, Bell System Tech. J. 37, 699 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

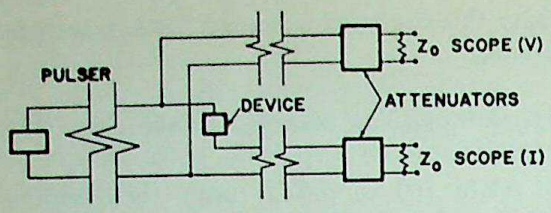


FIG. 2. Pulse measurement circuit.

precision decade resistance box to sense the current. In many instances, however, the dc data were obtained using an oscilloscope as an X-Y plotter. This procedure speeded up the data taking process considerably while still allowing accuracy within the range required. In some cases the device characteristics were observed on a curve tracer when high accuracy was not needed.

For currents greater than 10^{-3} A where heating would prevent dc measurements, the steady-state current-voltage characteristics were obtained using repetitive submicrosecond pulses. The pulse measurement setup, as shown in Fig. 2, differed slightly from the dc setup since to prevent reflections the transmission lines were matched with their characteristic impedance Z_0 at all terminations. The pulse duration in these measurements varied between 100 and 300 nsec. The repetition rate was approximately 150–350 pulses per second. A fast rise-time oscilloscope (Tektronix 555) was used to observe the pulses in all cases.

The single-shot pulse measurements, used to approximate the trap-free SCLC, were taken with the same circuit as shown in Fig. 2. A single-shot measurement was made by pulsing the device once after it had been thermally cycled from 300 to 4.2°K. The oscilloscope traces corresponding to the single-shot voltage and current pulses were recorded on high speed film, providing records from which the data could later be taken and analyzed.

EXPERIMENTAL CURRENT-VOLTAGE CHARACTERISTICS

The experimental steady-state current-voltage characteristics as obtained at 4.2°K for p_+pp_+ silicon devices exhibited the structure one would expect for SCLC in an insulator with deep trapping levels. In Fig. 3 the steady-state characteristic has been plotted for device II. This characteristic, which is typical of all the devices tested, clearly shows the presence of a traps-filled voltage, at which point the current increases steeply with voltage. When the traps-filled voltage is measured for this device and for the others constructed from the same material, the values tabulated in Table I are obtained. If the trap density were uniform, as was previously assumed above, a d^2 dependence would be observed in V_{TFL}. By plotting $\log V_{TFL}$ versus $\log d$, where d is the minimum thickness, a straight line of slope two would connect the measurements for the various devices for this dependence. In Fig. 4 this plot is shown including the possible errors in device thickness

TABLE I. Traps-filled voltage values and the corresponding minimum thicknesses for several devices.

Device	Minimum thickness (cm)	V_{TFL} (V)
O	7.7×10^{-3}	31
II	6.8×10^{-3}	25
III	8.2×10^{-3}	36
IV	9.0×10^{-3}	44
VII	1.1×10^{-2}	57
VIII	7.2×10^{-3}	25
IX	8.2×10^{-3}	32.5

measurements ($\pm5\,\mu$). The experimental points lie very close to the predicted slope-two line. The traps-filled voltage value obtained for each device allows a determination of the trap density in the device. When this is done for all the devices in Table I, a mean trap density of 6.1×10^{12} cm^{-3} is obtained. The maximum deviation from this value in the devices measured was approximately 10%.

From the steady-state curve in Fig. 3 the form of the current-voltage characteristic in the Child's law region can be determined. However, the limited data in this region make the determination, with any degree of accuracy, rather difficult. To decrease the possible error associated with fitting the Child's law region, single-shot pulse measurements can be made, to approximate the trap-free characteristic. These measurements, as taken for the device in Fig. 3 are plotted as the boxed data points on the same curve. The agreement between the steady-state and single-shot curves in this figure indicates the usefulness of the single-shot measurements in obtaining an extension of the Child's law region.

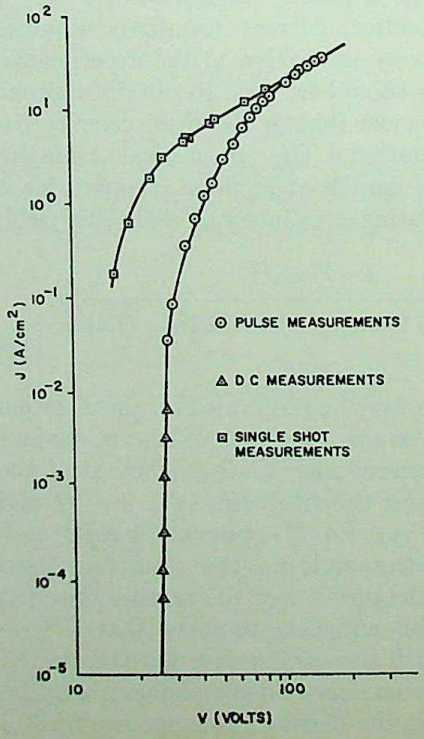


FIG. 3. Steady-state and single-shot pulse measurements for device II.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

By measuring the slope m of the characteristic in the extended Child's law region one can obtain from Eq. (4) $m=(2-\alpha)$. For the devices analyzed a mean value of 0.64 was obtained for α.

From the value of K_0 and the slope of the characteristic in the Child's law region, the field-dependent carrier mobility at low temperatures can be determined. In the range of applied fields 3×10^3 V/cm$<E<3\times10^4$ V/cm the relationship between mobility and electric field was found to be

$$\mu_P(E)\approx(2.6\pm1.4)\times10^5E^{-(0.64\pm0.04)}\ \text{cm}^2/\text{Vsec}. \quad (18)$$

The steeply rising region of the steady-state characteristic following V_{TFL} can be obtained analytically for the two cases of constant mobility or constant drift velocity. However, for the more general situation where the mobility is field-dependent but the drift velocity not saturated, no analytical solution can be obtained. To obtain the characteristic for this case a numerical solution would be necessary. The effort required to obtain this solution would not be justifiable since little additional information would result. However, the basic fact about the steeply rising region is that any model, which assumes the onset of SCLC after a traps-filled voltage, predicts an extremely steep characteristic regardless of the applicable $\mu(E)$ function. This prediction is most definitely verified by the previous experimental characteristics.

CAPTURE PROBABILITY AND CAPTURE CROSS-SECTION MEASUREMENTS

The mean capture probability,[5] $\bar{B}$, can be obtained from the same single-shot pulse measurements that were utilized to obtain the extended Child's law region of Fig. 3. Two pulse recordings, typical of those employed to obtain $\bar{B}$, are shown in Fig. 5 as photographed from the oscilloscope display.

There are two ways in which one can obtain $\bar{B}$ from the single-shot current response. One can (a) match the theoretical transient characteristic [Eq. (17)] to the experimental pulse, or (b) estimate the initial time constant of the decaying current pulse and compare it to the theoretical initial time constant.

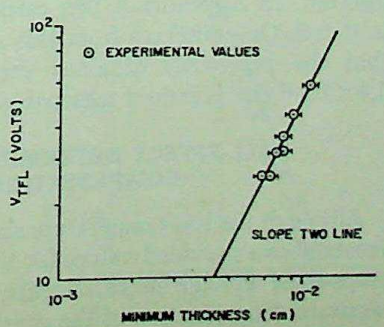


FIG. 4. Thickness dependence of V_{TFL}.

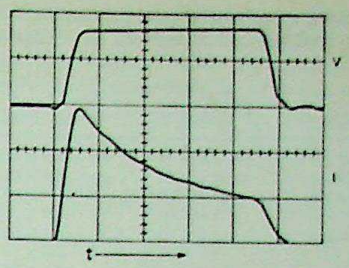


FIG. 5. Typical voltage and current transients obtained by single-shot pulse measurements. Scales: $t=40\times10^{-9}$ sec/major division: $V=25$ V/major division; $I=20$ mA/major division.

Equation (17) is rewritten for clarity as

$$j(t)=J_0\frac{1-V_{TFL}/V}{1-(V_{TFL}/V)\exp\bar{B}N_d(1-V/V_{TFL})t}.$$

It is seen to contain the parameters J_0 and $\bar{B}$ in addition to the previously determined constants N_d and V_{TFL}.

Some uncertainty exists as to where $t=0$ should be situated on the experimental pulse in order to cause it to correspond to the analytical characteristic of Eq. (17). Due to the accumulated rise time of the generator, oscilloscope, and measuring circuit there exists an uncertainty of approximately 20 nsec. Two alternative methods of matching the theoretical and experimental characteristics present themselves. One can assume

(a) that the beginning of the current pulse, $i(t)$, through the device corresponds to approximately 20×10^{-9} sec before the peak of the $i(t)$ pulse, or

(b) that negligible trapping occurs prior to the peak of the $i(t)$ pulse in which case this point would correspond to $t=0$ in the theoretical characteristic.

Case (a) involved matching the theoretical and experimental characteristics at two points to allow determination of both J_0 and $\bar{B}$ for each pulse. This technique was therefore called a "two-point fit." Case (b) involved matching the characteristics at only one point since J_0 was known. This case was called a "one-point fit." In Fig. 6 the experimental pulse responses of two typical devices are plotted as circled points. Both one- and two-point fits were performed for these units and the results are also shown. It was observed that for these and other pulses the one-point fit tended to more closely approximate the experimental results. This would indicate that no appreciable trap filling occurred for most pulses before the peak of the experimental characteristic.

The second general technique for calculating the mean capture probability from the experimental data was to relate the initial time constant τ_{TO} for decay of the current pulse to $\bar{B}$. Since at $t=0$ the density of trapping centers is approximately equal to N_D, one obtains that

$$B=1/\tau_{TO}N_D, \quad (19)$$

where τ_{TO} may be obtained by estimating the initial

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Capture probability and capture cross section values for negatively ionized boron centers in compensated silicon at 4.2°K.

Device	Method	$\bar{B}$ (cm³/sec)	$\sigma(\mathcal{E}_d)$ (cm²)	$\bar{E}$ (V/cm)
II	Initial slope	1.0×10^{-6}	2.0×10^{-13}	3.5×10^{3}
II	Initial slope	2.2×10^{-6}	4.0×10^{-13}	4.7×10^{3}
III	Initial slope	3.6×10^{-6}	6.9×10^{-13}	4.1×10^{3}
III	Initial slope	2.0×10^{-6}	4.9×10^{-13}	2.1×10^{3}
III	Initial slope	2.8×10^{-6}	6.1×10^{-13}	2.9×10^{3}
III	2-Point fit	3.2×10^{-6}	5.0×10^{-13}	7.1×10^{3}
III	1-Point fit	2.05×10^{-6}	3.4×10^{-13}	7.1×10^{3}
III	2-Point fit	1.6×10^{-5}	2.9×10^{-12}	4.8×10^{3}
III	1-Point fit	3.0×10^{-6}	5.5×10^{-13}	4.8×10^{3}
VII	Initial slope	4.0×10^{-6}	7.3×10^{-13}	4.9×10^{3}
VII	2-Point fit	6.0×10^{-6}	1.1×10^{-12}	4.9×10^{3}
IX	Initial slope	$(1.75\text{–}3.5)\times10^{-6}$	$(3.6\text{–}7.3)\times10^{-13}$	3.3×10^{3}
IX	Initial slope	1.5×10^{-6}	3.3×10^{-13}	2.7×10^{3}
IX	2-Point fit	3.8×10^{-6}	7.2×10^{-13}	4.4×10^{3}
IX	1-Point fit	2.2×10^{-6}	4.2×10^{-13}	4.4×10^{3}
IX	1-Point fit	1.4×10^{-6}	2.1×10^{-13}	8.3×10^{3}
IX	1-Point fit	1.4×10^{-6}	2.3×10^{-13}	6.6×10^{3}

slope of $i(t)$ on a semilogarithmic plot. For this reason this technique was called the "initial slope" method.

In Table II the values of $\bar{B}$ obtained by these three methods for several devices have been tabulated. The average electric field (V/d) is shown for each measurement. Because of the large scatter of the experimental results, no field dependence could be detected in the B measurements. By averaging the values for each of the techniques one obtains: (a) $\bar{B}=2.4\times10^{-6}$ from initial slope measurements, (b) $\bar{B}=2.0\times10^{-6}$ from one-point fitting of the theoretical characteristic, and (c) $\bar{B}=7.2\times10^{-6}$ from two-point fitting of the theoretical characteristic.

Since the value of $\bar{B}$ given in (c) does not agree well with (a) or (b) the value of $\bar{B}=2.2\times10^{-6}$ cm³/sec is chosen as the most probable.

In order to calculate a meaningful capture cross section for the trapping levels from the $\bar{B}$ value given above, one must know not only the energy distribution of the carriers but also the energy dependence of the capture cross section since $\bar{B}$ is given by Eq. (7). However, in the high-field case certain assumptions about the energy distribution of the carriers can be made which simplify the relationship between $\bar{B}$ and σ. For this case, because of the heating effect of the electric field on the carriers, optical mode scattering is the dominant energy-loss mechanism. If one assumes that the component of the carrier energy contributed by the field between optical collisions is much greater than the carrier energy remaining after a collision, the average carrier energy will be given by

$$\mathcal{E}_d=(1/2)m^*v_d^2, \tag{20}$$

where v_d is the drift velocity. If one also assumes that few carriers have energy is excess of $\mathcal{E}_R$, Eq. (7) simplifies to

$$\bar{B}=\sigma(\mathcal{E}_d)(2\mathcal{E}_d/m^*)^{1/2}. \tag{21}$$

Replacing the quantity

$$(2\mathcal{E}_d/m^*)^{1/2}$$

by its equivalent

$$K_0E^{1-\alpha},$$

one obtains from (21) the simple expression

$$\sigma(\mathcal{E}_d)=\bar{B}/K_0E^{1-\alpha}. \tag{22}$$

The values of $\sigma(\mathcal{E}_d)$ calculated by applying Eq. (22) to the values of $\bar{B}$ in Table II have been included in this table. The mean value of σ as taken from the initial slope and one point fit calculations is

$$\bar{\sigma}\approx4.4\times10^{13}\text{ cm}^2.$$

This value applies for a mean electric field of 4.6×10^3 V/cm.

The above experimental capture cross section can be compared to the prediction of the giant trap model by calculating the theoretical value of the energy-dependent capture cross section via relationships derived by Lax.[5] For the case of trapping via optical-phonon-emission Lax's theory predicts values for the capture cross section in the range

$$1.4\times10^{-13}\text{ cm}^2<\sigma<5.6\times10^{-13}\text{ cm}^2,$$

when constants appropriate to silicon at 4.2°K are inserted. In Appendix A this calculation is performed in detail. One observes from this range of values of σ that the previously obtained experimental value of 4.4×10^{-13} cm² is indeed reasonable.

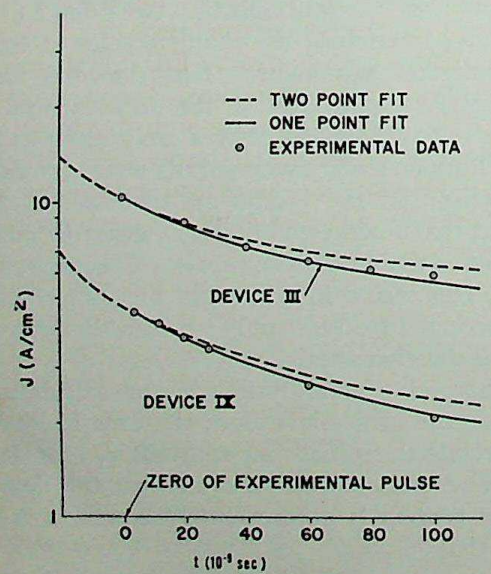


FIG. 6. Theoretical and experimental transient characteristics for devices III and IX.

HALL-EFFECT DETERMINATION OF COMPENSATION

Although the low-temperature single injection experiments above provided values for the compensation via traps-filled voltage measurements, it was felt that a separate determination of the compensation by a more

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

conventional technique was necessary for corroboration. For this purpose Hall-effect measurements were performed over the temperature range $25°K < T < 77°K$, whereby one could obtain both the majority dopant depth and degree of compensation.

The temperature-dependent hole concentration in p-type silicon at low temperatures is known to be given by the expression

$$p = \left(\frac{2\pi m^* T}{h^2}\right)^{3/2} \frac{(N_A - N_D)}{2N_D} \exp(-\mathcal{E}_A/kT), \quad (23)$$

where $\mathcal{E}_A$ is the acceptor activation energy.

To obtain the compensation N_D from the Hall data one observes that it is necessary to plot the quantity $\ln(pT^{-3/2})$ versus $1/T$. Over the region of applicability of Eq. (23) this plot yields a straight line of slope $-\mathcal{E}_A/k$ from which one obtains the acceptor activation energy. Furthermore, if the linear region of the plot is extrapolated to the $1/T=0$ axis, the intercept allows a determination of the quantity $(N_A-N_D)/N_D$. Since (N_A-N_D) can be obtained from Hall measurements in the exhaustion region, a measurement of the intercept readily yields the compensation N_D.

For three Hall samples this plot is shown in Fig. 7. Device I–II was constructed of material diffused simultaneously with that of the single injection devices previously considered. Device III–II and III–III were fabricated from the results of a separate diffusion performed under near identical circumstances on the same starting material. The results for these three samples are tabulated in Table III. One observes that the results for devices III–II and III–III are quite similar. However, the degree of compensation in these two devices differs by a factor of 2 from that measured for device I–II. This result supports the hypothesis that most of the compensation present in the devices was created during the diffusion step. The compensation observed in device I–II (6.6×10^{12} cm^{-3}) was quite close to the average value of 6.1×10^{12} cm^{-3} obtained from the previous V_{TFL} measurements. This degree of agreement between the two values is quite good, especially since finding N_D in the Hall experiment involves extrapolating the $pT^{-3/2}$ curve over a region as extensive as the straight line region itself.

TABLE III. Impurity concentrations and acceptor depth as determined by Hall-effect measurements.

Device	Type	E_D (eV)	(N_A-N_D) (cm^{-3})	N_D (cm^{-3})	N_A (cm^{-3})
I–II	p	47×10^{-3}	1.7×10^{12}	6.6×10^{12}	8.3×10^{12}
III–II	p	47×10^{-3}	5.4×10^{12}	3.9×10^{12}	9.3×10^{12}
III–III	p	47×10^{-3}	5.6×10^{12}	3.8×10^{12}	9.4×10^{12}

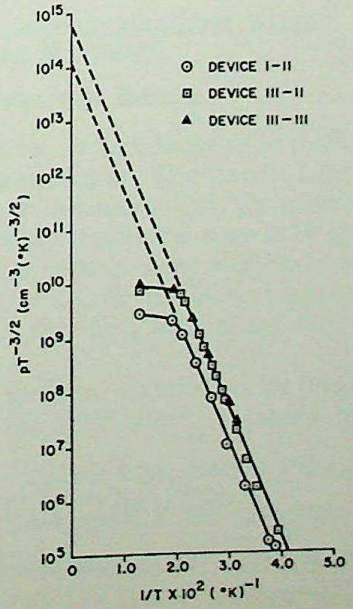


FIG. 7. $pT^{-3/2}$ versus $1/T\times10^2$ for Hall devices I–II, III–II and III–III.

The activation energy $\mathcal{E}_A$ of the acceptor center, corresponding to the trapping depth at 4.2°K, was approximately 0.047 eV for all devices. This value is known to correspond to boron in silicon.[11]

DISCUSSION AND CONCLUSIONS

In considering the model for p_+pp_+ devices at liquid-helium temperature a threshold voltage for injection is found to exist. This threshold is due to the barrier created at the p_+p junction by holes trapped near the junction in the equilibrium situation. The exact form of the Child's law region at low temperatures is dependent on the mobility-electric field relationship appropriate to the range of fields employed. For a mobility dependence as given by expression (3) the general Child's law relationship has the form of Eq. (4).

In order to relate the observed single-shot current pulses to the trapping capture probability, an approximate solution has been obtained for the current as a function of time under the conditions of slow trapping. This solution is as indicated by Eq. (17) if the current flow is assumed divergenceless. This assumption is strictly valid only for the case of drift-velocity saturation; therefore for $\alpha<1$ in expression (3) a slight disparity between the analytical expression for the current density and the actual characteristics is expected to exist.

The experimental results indicate that the SCLC model, modified to include the appropriate field-dependent mobility, applies for p-type silicon at 4.2°K. The steady-state current-voltage characteristics observed at this temperature correspond in form and magnitude to the predictions of the SCLS model for an insulator with a deep trap density of 6.1×10^{12} cm^{-3}. By single-shot pulse techniques the experimental Child's law regions could be extended to lower applied voltages.

[11] N. B. Hannay, *Semiconductors* (Reinhold Publishing Corporation, New York, 1960), p. 472.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Based on these extended power law regions the appropriate form of Child's law was found to be $j \sim V^{1.36}$. The mobility implied by the characteristic in this region was

$$\mu_P \simeq (2.6 \pm 1.4) 10^5 E^{-(0.64 \pm 0.04)} \text{ cm}^2/\text{Vsec}.$$

A trap density of 6.1×10^{12} cm^{-3} was obtained by trap-filled voltage measurements and it should correspond to the compensation in the material. This compensation was measured separately by a Hall-effect experiment over the temperature range $25°\text{K} < T < 77°\text{K}$ and found to be approximately 6.6×10^{12} cm^{-3}, which agrees reasonably well with the value above. Also from the Hall experiment a value for the acceptor level of 0.047 eV was obtained, which agrees reasonably well with the expected value for boron in silicon.

Experimental values for the capture probability and capture cross section were obtained by fitting the analytical characteristic for the current density to the observed pulses. The average experimental value for the mean capture probability was $\bar{B} = 2.2 \times 10^{-6}$ cm^3 sec^{-1}. If all the carriers are assumed to have the same velocity, this value could be converted to a corresponding one for capture cross section. The latter, as obtained by averaging the values obtained from the many experimental measurements of $\bar{B}$, was $\bar{\sigma} = 4.4 \times 10^{-13}$ cm^2. The theoretical value of σ, as predicted by the giant trap model, was calculated for the case of trapping via optical-phonon emission. A value of σ in the range $1.4 \times 10^{-13} < \sigma < 5.6 \times 10^{-13}$ cm^2 resulted. The agreement between experimental and theoretical values is again quite reasonable.

APPENDIX A

For the case of trapping via optical-phonon emission, Eq. (A1) relates σ and $\mathcal{E}$, the carrier energy. This equation is

$$\sigma(\mathcal{E}) = \frac{\pi^2}{8} \frac{w}{l_c} \frac{q^2}{\epsilon} \left(\frac{q^2}{\epsilon \hbar\omega} \right)^2 \times \frac{\lambda}{1-\exp(-\lambda)} \frac{P(\hbar\omega - \mathcal{E})}{(1-\mathcal{E}/\hbar\omega)^{5/2}} (1+a\bar{y})^{1/2}, \quad \text{(A1)}$$

where the various quantities are defined as: $\hbar\omega = \mathcal{E}_r$ the Raman energy (0.063 eV for silicon), $\lambda = \hbar\omega/kT$, $P(\hbar\omega - \mathcal{E})$ = the "sticking probability," $a = ((\hbar\omega - \mathcal{E})/\hbar\omega)$, $\bar{y} \approx 0.85$, l_c = the mean free path for acoustical scattering, and w = a constant which relates the strength of coupling of the carriers to the optical modes to the strength of coupling to acoustical modes.[12] After substituting the appropriate values of the above terms for silicon at 4.2°K into Eq. (A1) one obtains that

$$\sigma = 8.5 \times 10^{-16} w/l_c, \quad \text{(A2)}$$

when $E = 4.6 \times 10^3$ V/cm.

Choosing the pure acoustical scattering mobility at 4.2°K to be approximately 10^6 cm^2/Vsec for p-type silicon, as the results of Putley and Mitchell[6] would indicate, l_c would be approximately 6×10^{-4} cm. Therefore at 4.2°K

$$\sigma \approx 1.4 \times 10^{-13} (w) \text{ cm}^2. \quad \text{(A3)}$$

The value of w can be estimated based on the existing data for mobility versus temperature in the range $100°\text{K} < T < 300°\text{K}$.[6] It would appear that $1 < w < 4$ is applicable for p-type silicon.

For this range of w one finds that

$$1.4 \times 10^{-13} \text{ cm}^2 < \sigma < 5.6 \times 10^{-13} \text{ cm}^2.$$

[12] C. Herring, Bell System Tech. J. 34, 237 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Ultrasonic Wave Propagation in Doped *n*-Germanium and *p*-Silicon

WARREN P. MASON AND T. B. BATEMAN
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received 9 January 1964)

The effect of doping germanium with *n*-type material and silicon with *p*-type material is to increase the attenuation and decrease the elastic moduli. The decrease in the c_{44} elastic modulus for *n*-type germanium agrees fairly well with theoretical predictions. However, the modulus decrease in *p*-type silicon is much larger and varies with temperature much faster than predicted by any present theory. It is suggested that there is a temperature-induced change in hole population along the energy surfaces. The intervalley relaxation time, which determines the added attenuation, becomes independent of the doping for high dopings with the antimony time being about 100 times that for arsenic. This result indicates that the relative values are determined by the square of the triplet singlet separation which occurs near the impurity atoms. For *p*-type silicon the relaxation time at low temperatures increases very markedly indicating an activation-energy effect. The energies agree well with the energies measured by infrared techniques for the largest excited state orbits around the impurity atoms. The relaxation times measured at high temperatures indicate that the hole is transported $\frac{1}{3}$ cycle from one ⟨111⟩ position to the next.

I. INTRODUCTION

IT was first pointed out by Keyes[1] that the introduction of arsenic, phosphorus, or antimony into germanium—*n*-type doping—produces a decrease in the shear elastic modulus c_{44}. The result was shown to be due to the effect of the electronic part on the strain energy function and to involve only those electrons which change from one valley to another. The experimental results of Keyes and Bruner showed that the measured decrease was about 75% of the calculated decrease and varied with temperature as predicted by theory.

Pomerantz, Keyes, and Seiden[2] showed that the finite time required for electrons to go from one valley to another was a relaxation effect and should result in an attenuation of sound waves at a high frequency. Measurements at 9 kMc/sec did show[2] an absorption so high that it could not be quantitatively determined for any wave involving the c_{44} elastic modulus, while normal phonon-phonon attenuation resulted for any mode not involving this elastic constant.

II. EFFECT OF DOPING GERMANIUM WITH *n*-TYPE MATERIAL

A. Experimental Results

By going to 500 Mc/sec, or less, measurable results were obtained[3] for arsenic doping. The present paper extends these results to antimony and phosphorus doping. These data are of considerable interest in delineating the mechanisms causing changes of electrons from one valley to another and the relaxation times for these processes. Hence, they have been rather extensively investigated.

Figure 1 shows the increased attenuation for longitudinal waves along the ⟨110⟩ direction caused by doping germanium with 10^{18} and 3×10^{19} arsenic atoms per cc. The measurements were made at a frequency of 475 Mc/sec. The difference between the attenuation of the doped and pure crystal is practically independent of the temperature. The elastic modulus is determined by the sum

$$c=(c_{11}+c_{12}+2c_{44})/2. \quad (1)$$

Since it has been shown[1,3] that the only elastic constant that is affected by *n*-type doping of germanium is c_{44}, the attenuation can be ascribed to the results of a resolved shearing stress associated with the longitudinal wave. In fact the attenuation is caused by the finite relaxation time associated with the motion of electrons between the four valleys of the energy surface of *n*-type germanium. It satisfies the relaxation equation

$$A_{(\mathrm{Np./cm})}=(\Delta c_{44}/2\rho V^3)(\omega^2\tau), \quad (2)$$

valid when $\omega\tau\ll1$ and $(\omega/V)\lambda<1$, where λ is the mean free electron path. In this equation Δc_{44} is the change in modulus caused by the doping, ω is $2\pi f$, ρ the density, and V the velocity for the longitudinal wave. Changes in the c_{44} modulus for a doping of 3×10^{19} boron atoms per cc were measured by Bruner and Keyes[1] with the results shown by Fig. 2. This change was determined by the difference in velocities for pure and doped samples measured at 10 Mc/sec. The dashed curve shows the calculated value—from Eq. (13)—for the sample doped with 10^{18} arsenic atoms per cc. Measurements for antimony-doped germanium, shown by Fig. 4, give a very similar change.

The attenuation measurements of Fig. 1 and the velocity measurements of Fig. 2 can be used to determine the relaxation times, which are

$$\tau=4\times10^{-13}\ \mathrm{sec};\quad \tau=2.3\times10^{-13}\ \mathrm{sec}, \quad (3)$$

respectively, for dopings of 10^{18} and 3×10^{19} arsenic atoms per cc. These times are shown plotted on Fig. 3. On account of the small change in attenuation, this method cannot be used to investigate dopings less than 10^{18}

[1] R. Keyes, I.B.M. J. Res. Develop. **5**, 266 (1961); L. J. Bruner and R. W. Keyes, Phys. Rev. Letters **7**, 55 (1961).
[2] M. Pomerantz, R. W. Keyes, and P. E. Seiden, Phys. Rev. Letters **9**, 312 (1962).
[3] W. P. Mason and T. B. Bateman, Phys. Rev. Letters **10**, 151 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

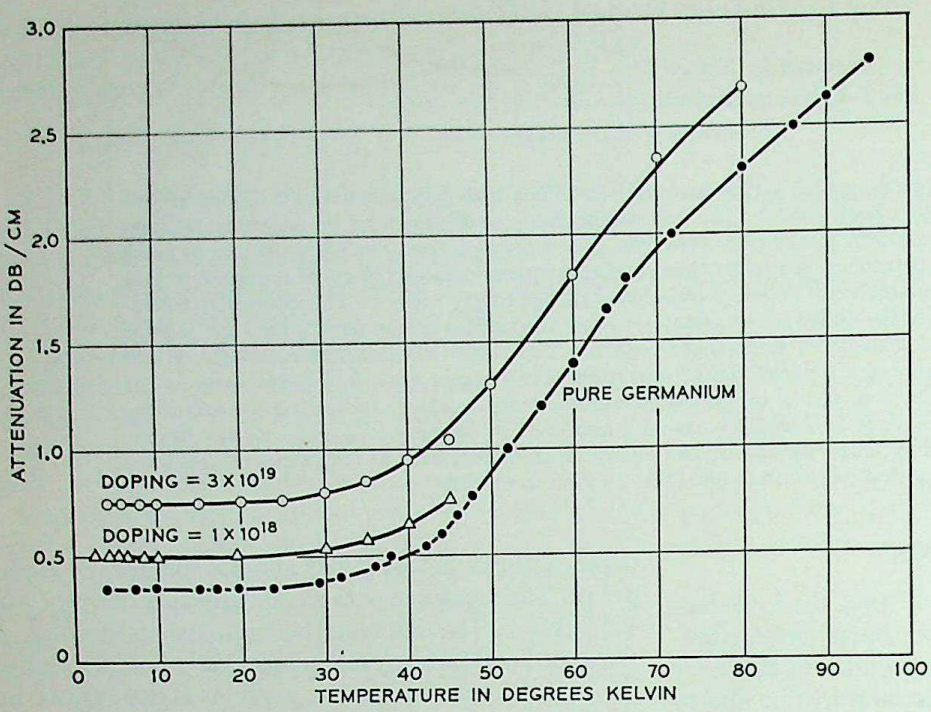


FIG. 1. Attentuation measurements of pure and arsenic-doped germanium. Longitudinal waves along ⟨110⟩, frequency 475 Mc/sec.

arsenic atoms per cc. However, another method applicable for lower dopings is the acousto-electric voltage method of Weinreich *et al.*[4] Figure 3 shows the relaxation times for a number of doping levels from 10^{14} to 10^{16} arsenic atoms per cc. At low temperatures the relaxation time is determined by impurity scattering and the solid line of Fig. 3 shows the values obtained as a function of doping using the 40°K values.[4] As the temperature increases, scattering occurs also by collisions with phonons and the dashed lines of Fig. 3 show the modification caused by this type of scattering for the temperatures shown.

The relaxation time for germanium doped lightly with arsenic impurities has been discussed by Weinreich *et al.*[4] as being due to a combination of scattering by neutral and ionized atoms. At low temperatures neutral scattering predominates and this gives a time inversely proportional to the impurity density. If we extrapolate this down to an impurity density of 10^{18}, the indicated relaxation time of 2×10^{-13} is less than the measured value of about 4×10^{-13} sec. For a density of 3×10^{19} the relaxation time approaches the value of 2.3×10^{-13} sec, which appears to be an intrinsic time for going from one valley to another.

An even stronger indication of this effect is furnished by antimony-doped germanium. Figure 4 shows the modulus and attenuation difference between a pure germanium sample and a sample doped with 1.5×10^{18}

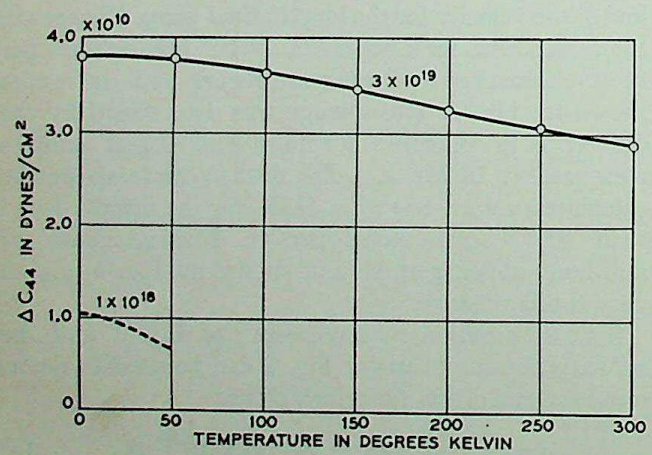


FIG. 2. Change in the elastic modulus c_{44} for arsenic-doped germanium. Measurements made at 10 Mc/sec. Dashed curve shows estimated value for a doping of 1×10^{18} arsenic atoms per cc (after Bruner and Keyes).

[4] G. Weinreich, T. M. Sanders, and H. G. White, Phys. Rev. **114**, 33 (1959).

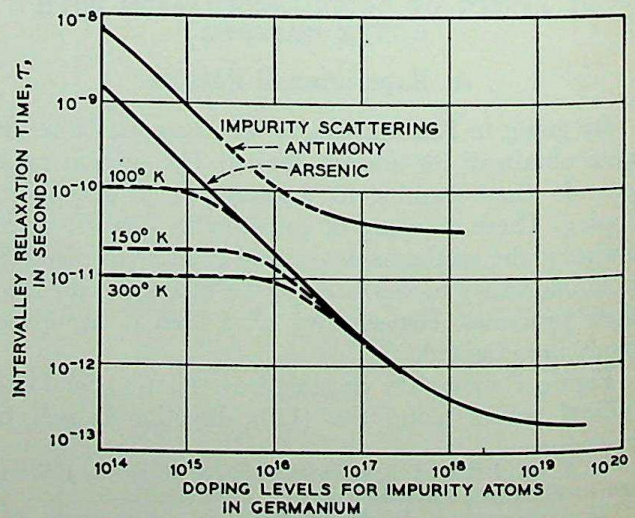


FIG. 3. Relaxation times for arsenic- and antimony-doped germanium as a function of doping. Dashed curves show effect of phonon scattering while solid curves are for impurity scattering.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

antimony atoms per cc. The attenuation is so high for this sample that a measuring frequency of 99 Mc/sec was used. As before the measurements were made along a ⟨110⟩ direction for longitudinal waves. The relaxation time calculated from these measurements is shown plotted by the dashed line of Fig. 4. At low temperatures impurity scattering is dominant and a value of 3.8×10^{-11} sec results. As the temperature is increased, phonon scattering occurs and the value approaches the same value found for arsenic doping at 300°K. The values for impurity doping by antimony are shown by Fig. 3. Another measurement was made for a doping of 10^{17} giving the value of the relaxation time, as shown. Measurements for lower doping have been made by Tell and Weinreich.[5] The results quoted indicate a relaxation time five times as large as that for antimony doping over a range from 10^{14} to 2.5×10^{15} atoms per cc. The resulting relaxation time for the entire range is shown by Fig. 3.

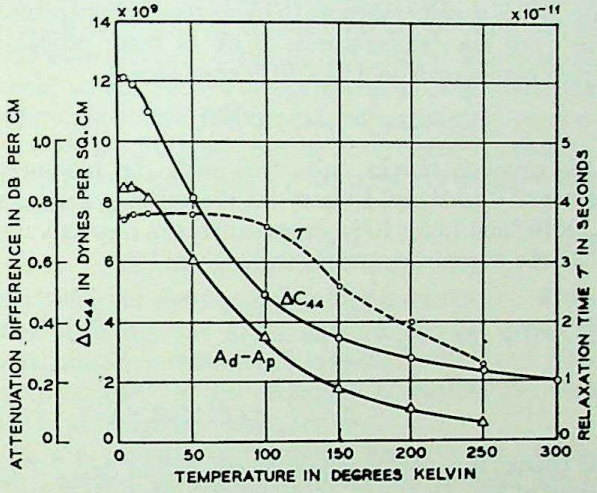


FIG. 4. Attenuation and modulus changes for germanium doped with 1.5×10^{18} antimony atoms per cc. Measurements made at 99 Mc/sec. Dashed line shows corresponding relaxation time.

Here the constancy of the intervalley relaxation time is confirmed over a wide doping range.

B. Theoretical Interpretation

Many experimental results for n-type germanium and silicon have been interpreted in terms of the many-valley model of the energy surfaces. Figure 5 shows the energy surfaces for silicon and it is seen that for the conduction band, one of the surfaces has a very low energy value in k space along the ⟨100⟩ axis. Electrons tend to congregate about these momentum values. Since there are six ⟨100⟩ directions, there are six energy minima or valleys. The six surfaces are shown by Fig. 6. The ovals indicate surfaces of constant energy. Similar considerations apply to the conduction band for germanium except that the minimum energy directions lie

[5] B. Tell and G. Weinreich, Bull. Am. Phys. Soc. 7, 546 (1962).

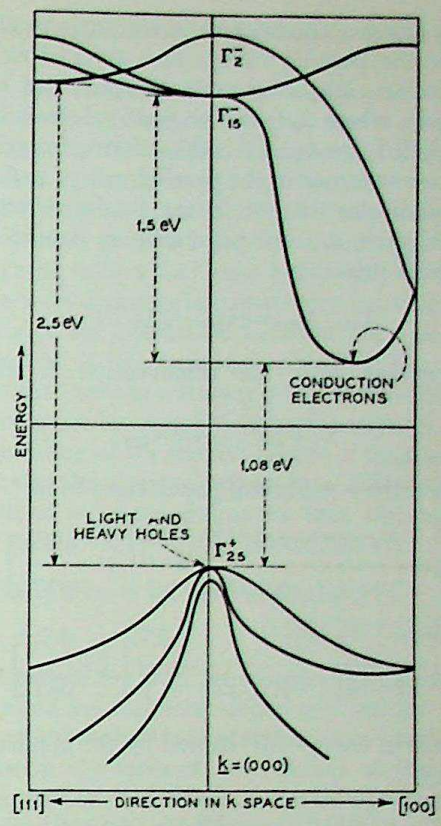


FIG. 5. Energy surfaces in valence and conduction band for silicon.

along ⟨111⟩ axis and occur at the edges of the Brillouin zone.

In the unstressed condition all the valleys have equal energies and electron populations. The effect of a

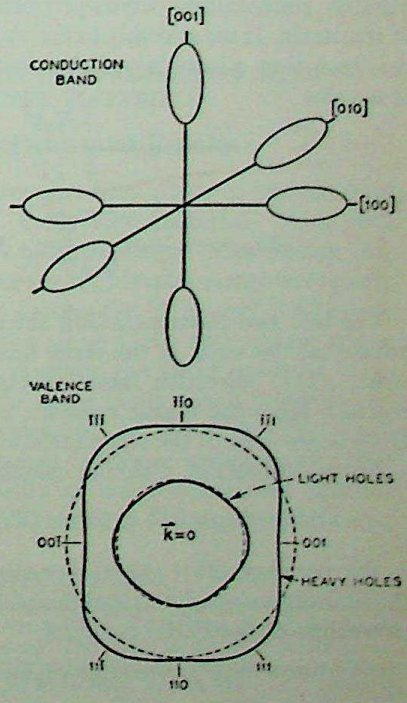


FIG. 6. Top figure: multivalley energy surfaces for n-silicon. Bottom figure: cross section of energy surfaces a constant distance below the origin.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

stress T_1 along a crystal axis in silicon is to raise the energy of the parallel valleys and lower those of the perpendicular valleys by amounts $\frac{2}{3}\Phi S_1$ and $-\frac{1}{3}\Phi S_1$, respectively, where S_1 is the longitudinal strain induced by the applied stress and Φ is the deformation potential. This causes electrons in the parallel valleys to flow into the perpendicular valleys. In the nondegenerate range the equilibrium ratio of population is determined by Boltzmann's principle

$$N_{100}/N_{010}=\exp(-\Phi S_1/kT). \tag{4}$$

This, together with the conservation of electrons given by

$$N_0=2N_{(100)}+4N_{010}, \tag{5}$$

determined the equilibrium populations to be

$$\begin{aligned} 2N_{100}&=\frac{N_0\exp(-\Phi S_1/kT)}{2+\exp(-\Phi S_1/kT)}\doteq\frac{N_0}{3}\left[1-\frac{2}{3}\frac{\Phi S_1}{kT}\right];\\ 4N_{010}&=\frac{2N_0}{2+\exp(-\Phi S_1/kT)}\doteq\frac{2N_0}{3}\left[1+\frac{\Phi S_1}{3kT}\right]. \end{aligned} \tag{6}$$

The change in energy ΔW caused by the applied stress is given by

$$\begin{aligned} \Delta W&=\frac{N_0}{3}\left[1-\frac{2}{3}\frac{\Phi S_1}{kT}\right]\frac{2}{3}\Phi S_1-\frac{2N_0}{3}\left[1+\frac{\Phi S_1}{3kT}\right]\frac{\Phi S_1}{3}\\ &=-\frac{2N_0}{9}\frac{\Phi^2 S_1^2}{kT}. \end{aligned} \tag{7}$$

There is no change to first powers of the strain but a negative contribution occurs proportional to the square of the strain. If we add this to the elastic energy $\frac{1}{2}c_{11}S_1^2$, it is seen that a lowering of the elastic modulus occurs of a value

$$\begin{aligned} \Delta c_{11}&=-\frac{4}{9}\frac{N_0\Phi^2}{kT};\\ \text{also}\quad \Delta c_{12}&=-\frac{\Delta c_{11}}{2}=-\frac{2}{9}\frac{N_0\Phi^2}{kT};\quad \Delta c_{44}=0. \end{aligned} \tag{8}$$

The last two relations follow since a shearing strain affects all the wells in the same fashion while a stress along ⟨111⟩ direction also affects all the valleys equally. Since the elastic constant for this direction is $(c_{11}+2c_{12}+4c_{44})/3$, the second relation of (8) must hold.

For n germanium, a similar calculation[1] shows that

$$\Delta c_{11}=\Delta c_{12}=0;\quad \Delta c_{44}=-(N_0/9)(\Phi^2/kT) \tag{9}$$

in the nondegenerate range. For the degenerate range the limiting value of Δc_{44} is obtained by replacing kT by two-thirds of the Fermi energy or

$$\tfrac{2}{3}W_F=(\hbar^2/3m^*)(3\pi^2N)^{2/3}, \tag{10}$$

where $\hbar$ is Plank's constant h divided by 2π, N is the number of free electrons, which is equal to the number of impurity atoms in the degenerate range, and m^* the density of states mass

$$m^*=(m_l m_t^2)^{1/3}(n)^{2/3}, \tag{11}$$

where m_l is the longitudinal mass of a single valley, m_t the transverse mass, and n the number of valleys. Inserting the value in (9), the limiting value of Δc_{44} at low temperatures is given by

$$\Delta c_{44_0}=-\frac{4}{3}\left(\frac{4\pi}{3}\right)^{2/3}\frac{N^{1/3}\Phi^2(m_l m_t^2)^{1/3}}{h^2}. \tag{12}$$

Since $m_l=1.57m_0$; $m_t=0.082m_0$, where m_0 is the electron mass 9.1×10^{-28} g, this equation takes the form

$$\Delta c_{44_0}=-1.58\times10^{25}\Phi^2N^{1/3}. \tag{13}$$

From piezoresistance measurements and from other effects Φ is usually taken as 16 eV at room temperature. For very low temperatures 19 eV is more probable which corresponds to 3.04×10^{-11} ergs. Hence,

$$\Delta c_{44_0}=1.44\times10^4N^{1/3}. \tag{14}$$

This equation checks quite well with the measured values of Figs. 2 and 4 where the theoretical values are 4.7×10^{10} and 1.65×10^{10} dyn per square cm, respectively.

At the degeneracy temperature, given by

$$\begin{aligned} kT_D=W_F\quad\text{or}\quad T_D&=\frac{\hbar^2}{2m^*}\frac{(3\pi^2N)^{2/3}}{k}\\ &=7.61\times10^{-11}N^{2/3}, \end{aligned} \tag{15}$$

the change in modulus should be one-half of Δc_{44_0}.[1] The curves of Fig. 4 indicate a value of $T_D=80°$K compared to the calculated value of 99°K. Over a temperature range the ratio of Δc_{44} to Δc_{44_0} follows the relation[1] shown by Fig. 7. Using $T_D=80°$K for Fig. 4, the agreement with the measured value is excellent.

The attenuation A is determined by the product of Δc_{44} and the relaxation time τ which is the time required to go from one valley to another. Measured values are

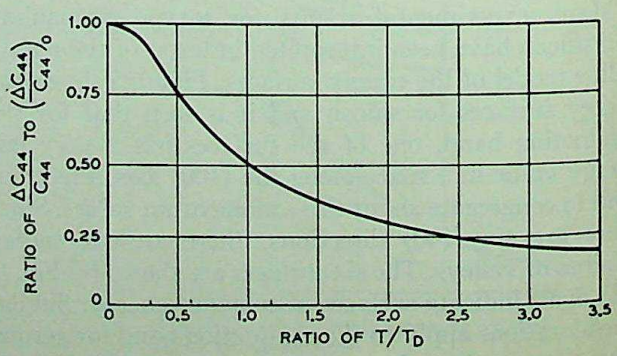


FIG. 7. Calculated ratio of $\Delta c_{44}/\Delta c_{44_0}$ as a function of the ratio of the temperature to the degeneracy temperature (after R. W. Keyes).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

shown by Fig. 3. A number of different mechanisms determine this time. For dopings from 10^{14} to 10^{16} and temperatures above 150°K, the scattering mechanism is the collision of electrons with phonons. The total relaxation time, as determined by the conductivity equation, is given by

$$\tau = m^*\sigma/Ne^2 = \mu m^*/e = 0.37\times 1.1\times 10^{-31}/1.6\times 10^{-19} \doteq 2.5\times 10^{-13}\ \text{sec}, \quad (16)$$

where σ is the conductivity in mho/meter, N the number of electrons per cubic meter, e the electronic charge of 1.6×10^{-19} C, μ the mobility in M^2/V sec, and m^* the effective drift mass in kilograms. Using the values for lightly doped samples, the indicated value is about 2.5×10^{-13} sec at 300°K. Hence, about one in every 40 collisions results in a change in valleys.

As the temperature becomes lower, or the doping higher, the intervalley relaxation time is controlled by scattering from impurities. For low dopings this process has been discussed by Weinreich *et al.*[4] as being due to a combination of scattering by neutral and ionized arsenic atoms. At low temperatures neutral scattering predominates and this gives a time inversely proportional to the impurity density. Scattering is probably due to elastic scattering by the impurity atoms which follows[6] a law

$$1/\tau = \pi N R^2 V, \quad (17)$$

where N is the density of impurity atoms per cc, R the effective radius, and V the electron velocity. At 40°K the indicated cross section R^2 is about 5×10^{-13} cm^2. This is attributed[4,7] to an exchange interaction between valleys although no values are given.

As the temperature increases some of the arsenic atoms become ionized and a component is found which varies inversely as the electron density. This component, which has a strong temperature variation, is attributed[4,7] to a capture of electrons in bound donor states followed by re-emission in another valley. As will be shown in the next section, this process has an activation energy and hence the time between capture and re-emission decreases as the temperature increases.

For densities above 10^{18}, the data of Figs. 1 and 4 show that there is no activation energy for the process and the data of Fig. 3 show that the relaxation time tends to reach a constant value independent of the doping level. All of the impurity atoms will be ionized since Fermi degeneracy[8] occurs in germanium at an impurity concentration between 10^{17} and 10^{18} impurity atoms per cc. The impurities for degeneracy[9] are in the order of 10^{17} for antimony and 5×10^{17} for arsenic doping. Hence, bound donor states do not exist for these dopings.

In the neighborhood of an impurity atom the four degenerate energy minima are split into a singlet and triplet states by the valley-orbit splitting. The separations caused by antimony, phosphorus, and arsenic are 0.57×10^{-3}, 2.9×10^{-3}, and 4.15×10^{-3} eV, respectively. One mechanism proposed for the interchange of electrons between valleys is a "giant trapping"[10] of electrons which involves a change in momentum between valleys by means of an interaction between the triplet and singlet states. If the electron is placed in the $1s$ state and simultaneously in a known valley, it will not be in a stationary state, but in a linear combination. At a later time, depending on the relative phases of the singlet and triplet state, the electron will be in a different valley. Weinreich *et al.*[4] have estimated that the relaxation time for such a valley change should be

$$\tau \geq (16/3)(\hbar/\delta E) = 8.4\times 10^{-13}\ \text{sec(arsenic)} = 6.14\times 10^{-12}\ \text{(Sb)}. \quad (18)$$

These values are too large for arsenic-doped germanium and too small for antimony-doped germanium.

Another derivation[11] indicates that the relaxation time should be in the ratio of the squares of the singlet-triplet separation and this is in better agreement with the experimental values. However, since this is essentially an elastic scattering process, the independence of the scattering time versus impurity density is still anomalous. The same is true for the total scattering time of Eq. (17). This time varies from 6 to 2×10^{-14} sec for the doping ranges from 10^{18} to 3×10^{19} impurity atoms per cc. The ratios of the intervalley to total scattering times are about 10 for arsenic and about 1000 for antimony.

III. EFFECT OF DOPING SILICON WITH p-TYPE MATERIALS

A. Experimental Results

Measurements have also been made of the effect of doping silicon with the p-type materials boron and gallium. The first measurements[3] were made for 2.5×10^{18} boron atoms per cc. Measurements were made for longitudinal waves along the ⟨110⟩ and ⟨100⟩ directions and shear waves along ⟨110⟩ with polarizations along ⟨001⟩—which involves the shear modulus c_{44}—and with polarization along ⟨1$\bar{1}$0⟩ which involves the $(c_{11}-c_{12})/2$ elastic modulus. As shown by Fig. 8(a), curve A, the attenuation for shear waves along ⟨100⟩ increases very markedly at low temperatures and becomes so high below 6°K and for a 475 Mc/sec frequency,

[6] A. H. Wilson, *The Theory of Metals* (Cambridge University Press, New York, 1958), p. 10.

[7] P. J. Price and R. L. Hartman (to be published) have recently given a derivation which accounts for the scattering of ionized arsenic atoms in germanium.

[8] F. J. Morin and J. P. Maita, Phys. Rev. 96, 28 (1954).

[9] H. Fritzche, Phys. Rev. 125, 1552 (1962).

[10] M. Lax, Phys. Rev. 119, 1502 (1960).

[11] P. J. Price, J. Appl. Phys. 31, 849 (1960); measurements have recently been made for one value (1.5×10^{18}) of doping by phosphorus atoms in germanium. The value of 8×10^{-13} sec is intermediate between As and Sb and agrees well with the square of the singlet-triplet separation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

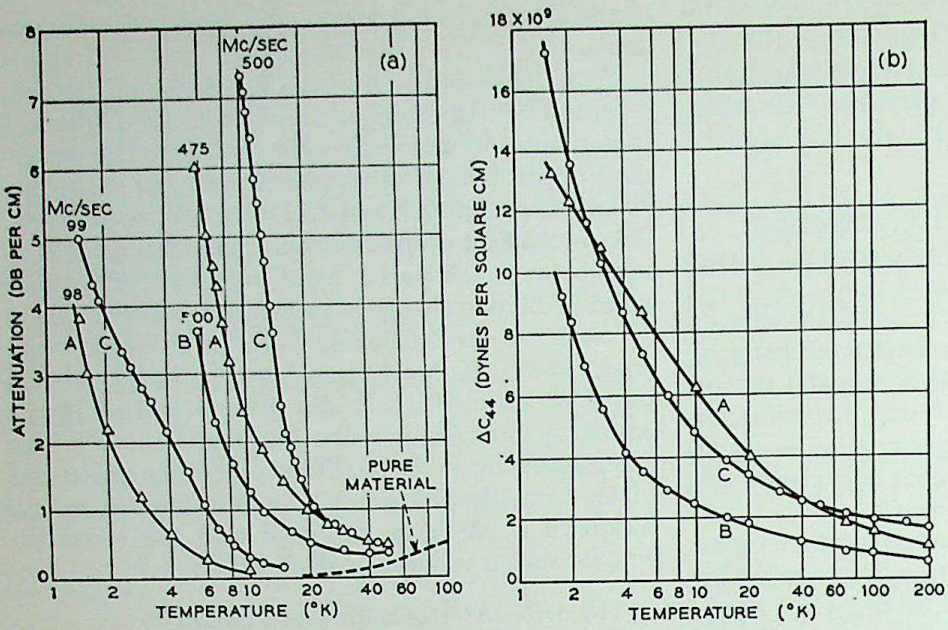


FIG. 8. (a) Measured attenuation as a function of temperature for A (2.5×10^{18} boron atoms per cc), B (5×10^{17} boron atoms per cc), C (1.5×10^{18} gallium atoms per cc). (b) Measured changes in the c_{44} elastic constants for the same conditions.

that it could not be measured. The frequency was shifted to 98 Mc/sec and the curve was measured from 1.5 to 15°K. In the overlap region, it is seen that the attenuation is square law as indicated by the relaxation Eq. (2).

Velocity measurements for pure and doped silicon were made by the pulse overlap method,[12] which is capable of good precision. Figure 9 shows the measured velocities for pure and doped samples. From these measurements, the change in modulus Δc_{44} can be determined and it is shown plotted by Fig. 8(b), curve A. The relaxation time from the measured attenuation and modulus differences can be calculated by using Eq. (2). The result is shown plotted on Fig. 10, curve A. For temperatures above 8°K the relaxation time satisfies the equation

$$\tau_1 = 1.3\times10^{-12}e^{13/T}, \qquad (19)$$

indicating an activation energy δE of 0.00113 eV. At lower temperatures the curve bends off and approaches

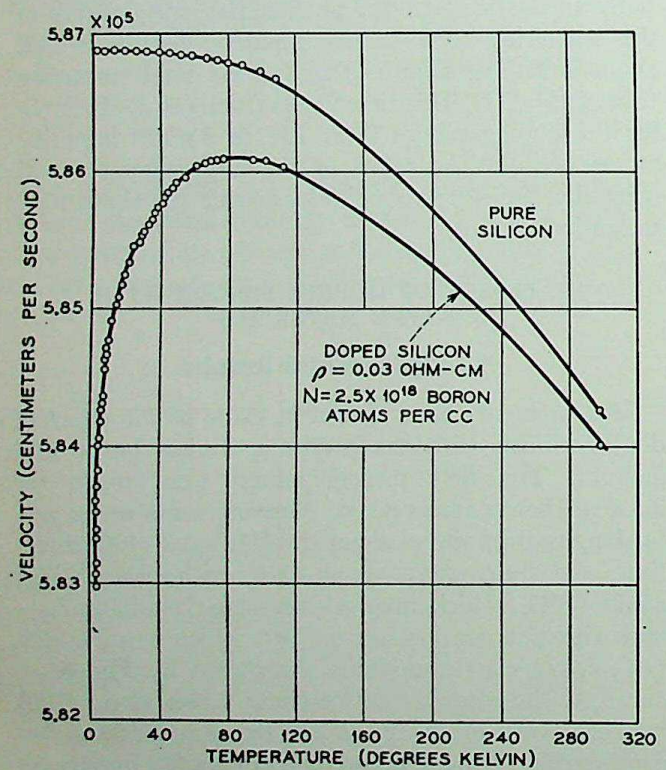


FIG. 9. Measured shear wave velocities along ⟨100⟩ direction for pure silicon and silicon doped with 2.5×10^{18} boron atoms per cc. Measurements made at 20 Mc/sec.

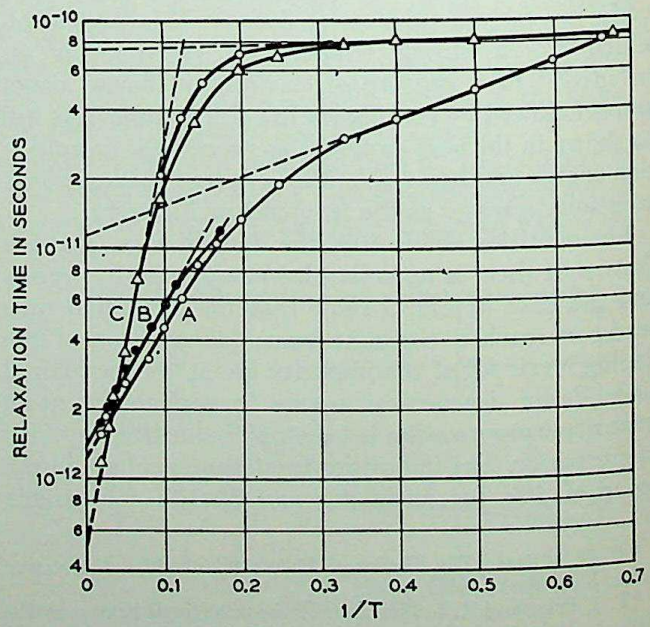


FIG. 10. Relaxation times for p-doped silicon plotted against $1/T$. A (2.5×10^{18} boron atoms per cc); B (5×10^{17} boron atoms per cc); C (1.5×10^{18} gallium atoms per cc).

[12] H. J. McSkimin and P. Andreatch, J. Appl. Phys. 34, 609 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

another straight line having an equation

$$\tau_1' = 1.13\times10^{-11}e^{2.95/T};\quad \delta E = 2.56\times10^{-4}\text{ eV}. \quad (20)$$

Detailed measurements were also made for longitudinal waves along the ⟨100⟩ direction with the results shown by Fig. 11. Here an increase in both attenuation and modulus change occurred as the temperature decreased, but their ratio, which determines the relaxation time, is approximately constant with a value $\tau_2 = 4.2\times10^{-12}$ sec.

The shear wave measurements along ⟨110⟩ with polarization along ⟨1$\bar{1}$0⟩ and the longitudinal measurements along ⟨110⟩ confirmed the expected relation that

$$\Delta c_{12} = -\Delta c_{11}/2. \quad (21)$$

The attenuations for these modes satisfy, respectively, the equations

$$A = \frac{\frac{3}{4}\Delta c_{11}\omega^2\tau_2}{2\rho V_s^3};\quad A = \frac{[(\Delta c_{11}/4)\tau_2 + \Delta c_{44}\tau_1]\omega^2}{2\rho V_l^3}. \quad (22)$$

The relaxation time τ_1 for shear waves along ⟨100⟩ has an activation energy term. In order to determine the origin of this term, measurements were made for a range of boron doping from 5×10^{17} to 3×10^{19} atoms per cc. Figures 8(a) and (b), curves B, show the results for shear waves along ⟨100⟩ measured at 500 Mc/sec for the doping of 5×10^{17}. The relaxation time is shown plotted on Fig. 10, with an indicated activation energy of 0.00117 eV or nearly the same as for the higher doping. The curves of Figs. 12 and 13 show the attenuation and velocities for the dopings of 10^{19} and 3×10^{19} boron atoms per cc. The lower value still has an increase in attenuation as the temperature is lowered, but the higher doping value has an added attenuation which is independent of the temperature with no activation energy. The doping of 10^{19} satisfies the relaxation equation

$$\tau_1 = 2.1\times10^{-13}e^{4/T}, \quad (23)$$

while the relaxation time for the doping of 3×10^{19} is 10^{-13} seconds independent of the temperature.

In order to complete the interpretation, a measurement was made for a doping of 1.5×10^{18} gallium atoms per cc in a ⟨100⟩ silicon crystal. Measurements were made for shear waves in this direction at a frequency of 500 Mc/sec. The results for the attenuation and

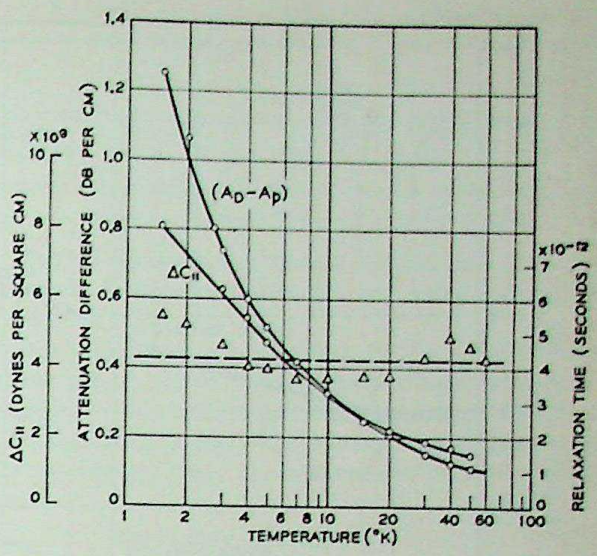


FIG. 11. Attenuation and modulus change Δc_{11} for longitudinal waves along the ⟨100⟩ axis. Dashed line shows an average relaxation time.

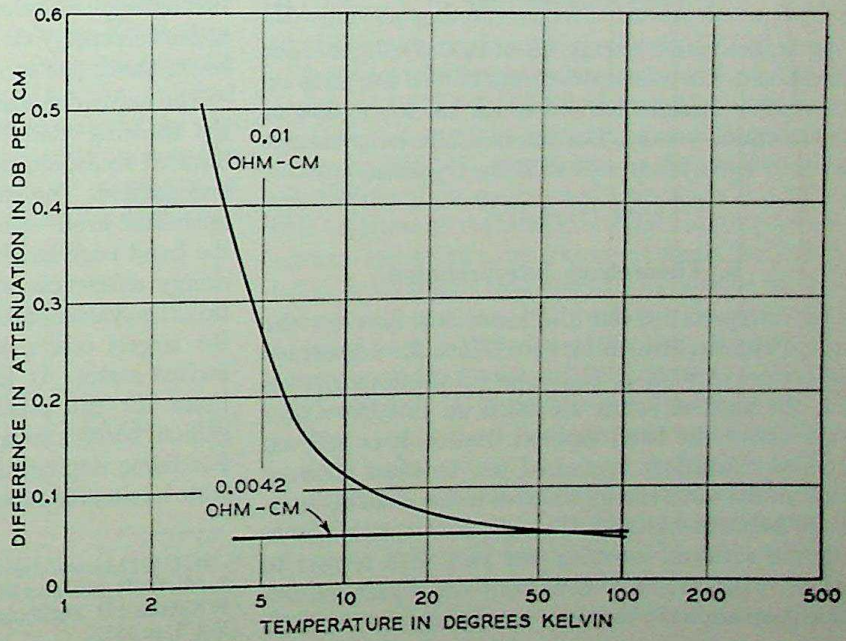


FIG. 12. Attenuation measured at 500 Mc/sec for boron-doped silicon as a function of the temperature. The curve labeled 0.01 Ω-cm is for a doping of 10^{19} boron atoms per cc while the one labelled 0.0042 is for 3×10^{19} boron atoms per cc.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

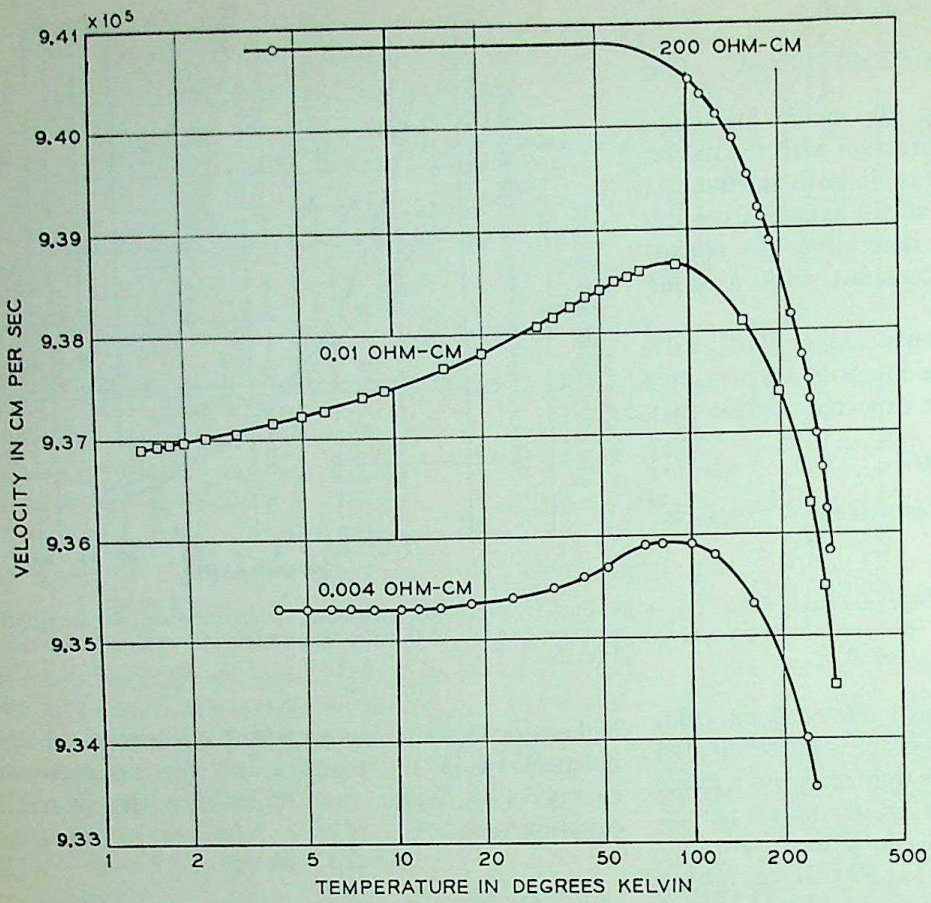


FIG. 13. Measured velocities for pure and doped silicon samples as described in Fig. 12.

modulus changes are shown by Figs. 8(a) and (b) by the curve labeled C. In the lower temperature range, measurements of the attenuation were made down to 1.5°K for a frequency of 99 Mc/sec. A plot of the relaxation time is shown by Fig. 10, curve C. The initial part of the relaxation time satisfies the equation

$$\tau_1 = 5\times 10^{-13} e^{40/T} \text{ sec}, \qquad (24)$$

giving an activation energy δE of 0.0035 eV. At lower temperatures the relation becomes curved but ends up with another straight line τ_1' which has a very low or zero activation energy. For intermediate temperatures the points agree fairly well with the equation

$$1/\tau = 1/\tau_1 + 1/\tau_1'. \qquad (25)$$

B. Theoretical Interpretation

The interpretation for the relaxation time for the p region appears straightforward. Since the degeneracy temperature is 50°K or higher for all the doping levels used, the valence bands are filled up to 0.004 eV or higher above the band minima. Hence, it is believed that the relaxations measured are between different points on the same energy surfaces rather than an interchange between surfaces. The effect of a strain is to warp the surfaces, lowering one part with respect to another. This causes a flow from the higher energy-momentum region to the lower energy regions. The holes change momentum directions near the impurity atoms by being captured in one of the excited states. This can happen with only a small number of collisions with phonons since the energies of the holes and the excited states are not very different. To leave the orbit, however, requires thermal energy to overcome the binding energy of the state, and this energy represent the activation energy measured. As the doping increases the activation energy decreases, and for a doping of 3×10^{19} boron atoms per cc it disappears entirely.

The activation energies for the relaxation time τ_1 of the shearing mode are in good agreement with the infrared measurements for the largest orbits in boron and gallium. The energy levels of the ground state—innermost orbit—and the excited states in relation to the band edge have been discussed by Kohn.[13] These energy differences were measured by infrared absorption. The values of 0.0012 and 0.0035 eV agree well with the largest orbits, lowest energies, for the measured excited states. At lower temperatures, the relaxation times for the primary modes for both boron and gallium become so large that other processes take over. For boron doping, the next process is another excited state having the constants of Eq. (20), while for gallium

[13] H. J. Hrostowski and R. H. Kaiser, Bull. Am. Phys. Soc. [2] 2, 66 (1957); see also W. Kohn in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press, Inc., New York, 1957), Vol. 5, p. 257.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

doping it appears to be a process without activation energy such as elastic scattering by the impurity atoms.

Another verification that the processes observed are the excited states comes from the high-temperature values of τ_1 and τ_1' for both boron and gallium doping. The frequency with which an orbit is traversed is given by

$$f=\delta E/h \doteq 2.8\times10^{11} \text{ cycles/sec (boron)}$$
$$\doteq 8.4\times10^{11} \text{ cycles/sec (gallium)} \quad (26)$$

for the first—τ_1—process. The time of a cycle is the inverse of these values and the measured constants of 1.3×10^{-12} sec (boron) and 5×10^{-13} sec (gallium) indicate that the transfer of momentum at high temperatures takes place between points that are about one-third of a complete orbit apart. Since the shearing stress raises or lowers the energy surfaces along ⟨111⟩ directions, the closest equivalent positions are 109° apart or about one-third of a complete cycle. The second process τ_1' of Eq. (20) indicates that the hole traverses two-thirds of a much larger orbit before it transfers its momentum.

As the doping increases, the impurity atoms come so close together that bound orbits are no longer possible. In this region the only process left is the elastic scattering by the ionized impurities which follows the law of Eq. (17). If we use the measured relaxation time of 10^{-13} sec for the impurity doping of 3×10^{19} boron atoms per cc, and calculate the velocity from the equation

$$\tfrac{1}{2}m^*V^2=0.6W_F=\tfrac{3}{10}(\hbar^2/m^*)(3\pi^2N)^{2/3}, \quad (27)$$

one finds for the scattering radius R the value of 8.5×10^{-8} cm. Comparing the relaxation time 10^{-13} with the total relaxation time from the conductivity Eq. (16) it appears that about 1 out of 10 collisions results in a change of position on the surface.

The theoretical situation for the change in moduli due to doping for p-type silicon is far from satisfactory. Three theoretical derivations[1,14,15] have been given for the change in shearing modulus. The first solution[1] assumes spherical energy surfaces, as shown by the dashed lines of Fig. 6, and limits consideration to the heavy hole surface. The second solution[14] takes account of the curvature of the two surfaces. In the degenerate range these solutions reduce to the forms

$$\Delta c_{44_0}=-\frac{(8\pi/3)^{2/3}m^*d^2N^{1/3}}{5\hbar^2}$$
$$=\frac{-0.825m^*d^2N^{1/3}}{\hbar^2} \text{ (Keyes)}, \quad (28)$$

$$\Delta c_{44_0}=(-2.38m^*N^{1/3}d^2/\hbar^2) \text{ (Bir and Tursunov)}. \quad (29)$$

[14] G. L. Bir and A. Tursunov, Fiz. Tverd. Tela 4, 2625 (1962) [English transl.: Soviet Phys.—Solid State 4, 1925 (1963)].

[15] P. Csavinszky and N. G. Einspruch, Phys. Rev. 132, 2434 (1963).

TABLE I. Values of Δc_{44} at 1.5°K for various dopings.

	Boron doping				Gallium doping
Doping	3×10^{19}	10^{19}	2.5×10^{18}	5×10^{17}	1.5×10^{18}
$\Delta c_{44}\times10^{-9}$	17.9	12.4	13.5	10.5	17.2

Table I shows measurements for Δc_{44} at 1.5°K for different dopings for boron and one value for gallium doping.

In the degenerate region from 10^{19} to 3×10^{19} boron atoms per cc, the variation is proportional to the one-third power of the doping level in agreement with Eqs. (29) and (30). Taking m^* as $0.44m_0$ as an average value for the two surfaces, the deformation potential values are

$$d=17.2 \text{ eV (Keyes)}; \quad (30)$$
$$d=10.1 \text{ eV (Bir and Tursunov)}.$$

A third calculation has recently been made[15] which yields values close to the value given by Keyes. Csavinszky and Einspruch[15] have considered the spin-orbit coupled band in addition to the two degenerate bands but have assumed spherical symmetry of the bands. For doping up to 10^{19} boron atoms per cc, the two upper bands are sufficient for the calculation and the result obtained is

$$\Delta c_{44_0}=-\frac{1}{5}\left(\frac{8\pi}{3}\right)^{2/3}\frac{d^2}{\hbar^2}[m_{v_1}N_1^{1/3}+m_{v_2}N_2^{1/3}], \quad (33)$$

where $m_{v_1}=0.49m_0$ is the heavy hole mass and $m_{v_2}=0.16m_0$ is the light hole mass. Since

$$N_1/N_2=(m_{v_1}/m_{v_2})^{3/2}=5.35, \quad (34)$$

$N_1=0.85N_0$ and $N_2=0.15N_0$. Hence, for a doping of 10^{19} boron atoms per cc, the value of $d=15.4$ eV.

However, the larger values measured at the lower dopings at 1.5°K—for which asymptotic values have not been reached—and the large variation with temperature shown by Fig. 8(b) cast doubt on the validity of these calculations. It is conjectured that there is a temperature variation of the position of the holes on the energy surfaces and at low temperatures it seems likely that the heavy holes collect near the ⟨111⟩ positions while light holes collect near the ⟨100⟩ positions. As seen from Fig. 6, lower part, which shows the cross sections of the energy surfaces a constant distance below the origin, the heavy-hole ⟨111⟩ positions and the light-hole ⟨100⟩ positions are farthest from the origin and require electrons of higher momentum to reach the surface. Hence, if momentum is equalized in all directions, the surfaces near the origin should be filled up to a level higher than those far away and hence the ⟨111⟩ positions for the heavy-hole surface and the ⟨100⟩ positions for the light-hole surface should have the lower energies. At very

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

low temperatures, holes tend to congregate in these directions.

If this conjecture is correct, the change in modulus for the heavy-hole surface should follow that for germanium, as given by Eq. (12), while that for the light-hole surface should follow Eq. (8) with kT replaced by (10), for the low temperatures, giving

$$\Delta c_{11_0} = -\frac{16}{3}\left(\frac{\pi}{3}\right)^{2/3}\frac{m^* N^{1/3}\Phi^2}{h^2}. \quad (31)$$

If we take an average value of 15 eV for the shear deformation potential d, the indicated change in modulus from Eq. (12)—using $m^*=0.49m_0$ and $N=0.85N_0$, i.e., 2.12×10^{18} heavy holes—is 26.5×10^9 dyn/cm^2, which is not out of line with the values of Fig. 8(b) extrapolated to lower temperatures. For higher doping levels this effect disappears since the momentum differences for the various positions become smaller.

Another piece of evidence that the light holes congregate along ⟨100⟩ positions at low temperatures is furnished by the nearly constant relaxation time of Fig. 11 for longitudinal waves propagated along the ⟨100⟩ direction. The data of Fig. 12 show that p-silicon with boron becomes degenerate for a doping of 3×10^{19} atoms per cc. This is taken to mean that the impurities are near enough together to prevent any excited state orbits around the boron atoms. The smallest orbits will be executed by the heavy holes and hence they will reach degeneracy at a higher doping level than the light hole surface. With a mass ratio of 0.49 to 0.16 or 3.06, the radius will be this factor larger for the light-hole surface. Hence, this surface should become degenerate for a doping of

$$3\times10^{19}/(3.06)^3 \doteq 10^{18} \text{ boron atoms per cc.} \quad (32)$$

Therefore, the constant relaxation time of 4.2×10^{-12} sec for a sample doped with 2.5×10^{18} boron atoms per cc is a confirmation that, at low temperatures, a longitudinal stress along the ⟨100⟩ axis actuates mostly light holes.

If the relaxation time is due to scattering of light holes by impurities, calculations indicate that the scattering radius is about 5×10^{-8} cm, in good agreement with the value obtained for the degenerate heavy-hole surface.

PHYSICAL REVIEW VOLUME 134, NUMBER 5A 1 JUNE 1964

Distribution Functions for the Number of Distinct Sites Visited in a Random Walk on Cubic Lattices: Relation to Defect Annealing*

J. R. Beeler, Jr.
Nuclear Materials and Propulsion Operation, General Electric Company, Cincinnati, Ohio

(Received 17 January 1964)

Distribution functions for the number of distinct sites $S(n)$ visited by a point defect executing a symmetric random walk of n jumps on two- and three-dimensional lattices were computed using the Monte Carlo method. The square planar, simple cubic, bcc, and fcc lattices were treated. In three dimensions, the normal distribution appears to describe $S(n)$ for $n>10^4$ jumps and at 10^4 jumps the derivative $d\bar{S}(n)/dn$ of the average number, $\bar{S}(n)$, of distinct sites is within 0.5% of the value given by Vineyard's exact asymptotic solution. The defect annealing rate was computed using the $S(n)$ distribution in a simple example and this result compared with an analog Monte Carlo solution. The comparison indicated that fluctuations in the initial defect concentration must be considered in computing the initial annealing rate and the mobile defect concentration as a function of time. After 500 jumps the annealing rate, but not the concentration, can be closely approximated without accounting for fluctuations in the initial concentration.

1. INTRODUCTION

IN a monatomic crystal, the migration of a defect, such as a vacancy or an interstitial atom, proceeds as a symmetric random walk[1] on the crystal lattice. The rate at which mobile defects are annihilated or trapped at point sinks is proportional to the rate at which they encounter fresh sites which have not been visited previously. Damask and Dienes[2] will treat the physical side of this process in a forthcoming book. On the basis of a Monte Carlo study, Beeler and Delaney[3] concluded that the average number of distinct sites, $\bar{S}(n)$, visited by a point defect in either a symmetric or an asymmetric random walk of n jumps was of the form,

$$\bar{S}(n)=An^k \quad (1)$$

* Work supported by the U. S. Atomic Energy Commission.

[1] In this discussion a symmetric random walk is one wherein the jump probabilities for each possible jump direction are equal and constant. The vacancy random walk in an alloy is in general asymmetric because of its ordering energy, i.e., the jump probabilities for each possible direction are not equal and also depend upon the position of the vacancy.

[2] A. C. Damask and G. J. Dienes, "Point Defects in Metals" (to be published).

[3] J. R. Beeler, Jr., and J. A. Delaney, Phys. Rev. 130, 962 (1963); J. R. Beeler, Jr., U.S.A.F. Report ASD-TDR 63-215 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

for a two-dimensional lattice, and of the form

$$\bar{S}(n)=B+Cn \tag{2}$$

for a three-dimensional lattice, where A, B, C, and k are constants for a given lattice structure. Their computations pertained to vacancy migration in a binary alloy and describe the relation between $\bar{S}(n)$ and the rate of ordering (disordering) for $n\leq 10^4$ jumps. Migration on the square planar simple cubic, bcc, and fcc lattices was investigated. Vineyard[4] subsequently proved that the asymptotic form of $\bar{S}(n)$ for a three-dimensional symmetric random walk is

$$\lim_{n\to\infty}\bar{S}(n)=Cn. \tag{3}$$

The Monte Carlo results of Beeler and Delaney for $d\bar{S}(n)/dn=C$ on the interval $10^3\leq n\leq 10^4$, in the case of zero ordering energy, agree closely with the exact asymptotic results given by Vineyard's analysis, the Monte Carlo results being about 1% larger than the exact asymptotic values. Vineyard also showed that

$$\bar{S}(n)=(8n)^{1/2}/\pi \tag{4}$$

in one dimension.

In this communication we describe the distribution function $F(S;n)$ for $S(n)$, as given by Monte Carlo calculations, for the square planar (sp), simple cubic (sc), bcc and fcc lattices. $F(S;n)$ is the probability that a defect visits S or less distinct lattice sites in n jumps. Direct application of these distributions to defect annealing calculations is illustrated in a simple, yet physically significant numerical example wherein attention is called to the effect of fluctuations in the initial defect concentration upon the initial defect annealing rate. This effect first came to our attention in a comparison of the annealing rate predicted by $F(S;n)$, on the basis of the macroscopic (average) defect concentration, with that obtained directly in a separate analog Monte Carlo annealing calculation wherein fluctuation effects were automatically accounted for. The presence of fluctuations significantly increased the initial annealing rate relative to that obtained on the basis of the macroscopic concentration.

On the interval $1000<n<5000$, $F(S;n)$ appears to start an approach to the normal distribution. A χ^2 analysis indicated that $F(S;n)$ is very probably normal for $n>10^4$ in three-dimensional lattices, but gives no information on the convergence to a normal distribution in two dimensions. Up to $n=10^4$ the density function $f(S;n)=dF(S;n)/dS$ is definitely skewed to the right. This feature enhances the increased initial annealing rate contribution arising from concentration fluctuations. The indicated increase is possibly large enough to be detectable in high-temperature quenching experiments and should exert a dominant influence during the initial stage of defect annihilation and clustering in irradiated specimens.

[4] G. H. Vineyard, J. Math. Phys. 4, 1191 (1963).

2. COMPUTATIONAL METHOD

Distribution functions $F(S;n)$ and density functions $f(S;n)$ for each lattice were estimated by running 1000 independent random walk histories of 10^4 jumps each using an IBM-7090 computer. The mean ($\bar{S}$), mode (S^*), median (S_m) and standard deviation (σ) were computed at 10-jump intervals up to 100 jumps, 100-jump intervals up to 10^3 jumps, and 1000-jump intervals up to 10^4 jumps. In each instance an expected tendency toward the normal distribution was monitored by performing an χ^2 test for 21 deg of freedom[5] and computing the coefficients of skewness (γ_1) and excess (γ_2).[6] $\bar{S}(n)$ and σ were computed using the first 500 histories, the second 500 histories and all 1000 histories. The differences between the results given by each sample set were found to be negligible and on this basis we decided that the 1000-history sample size was sufficiently large to form a basis for annealing calculations. In this regard, it is interesting to note that the difference between the present $\bar{S}$ results, based on a least-squares fit to data from 1000 histories, and the original calculations of Beeler and Delaney, based on a least-squares fit to the data from only 10 histories, is less than 0.5%.

3. DISTRIBUTION FUNCTION RESULTS

Selected plots of $F(S;n)$ for $n=10$, 20, 50, 100, 1000, and 10 000 jumps are given in Figs. 1–4 in terms of the ratio $S(n)/n$. This mode of presentation makes it easier to see how $F(S;n)$ changes as n increases, than if the curves were plotted versus $S(n)$. Let $S_1(n)$ be the largest S value such that $F(S;n)=0$ and $S_2(n)$ the smallest S wave such that $F(S;n)=1$. The range of the distribution is defined as $\Delta S(n)=S_2(n)-S_1(n)$. In all instances $\Delta S(n)/n$, $\bar{S}(n)/n$, $S_m(n)/n$, and $S^*(n)/n$ decrease monotonically as n increases. $S_1(n)/n$ remains more or less constant in the two-dimensional case, the diminution in $\Delta S/n$ arising essentially from a monotonic decrease in $S_2(n)/n$. In the three-dimensional

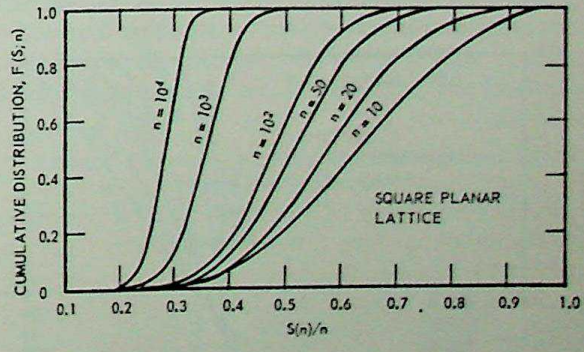


FIG. 1. Distribution function for square planar lattice.

[5] D. R. Evans, *The Atomic Nucleus* (McGraw-Hill Book Company, Inc., New York, 1955); see pp. 774–777.

[6] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1954); see p. 183.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

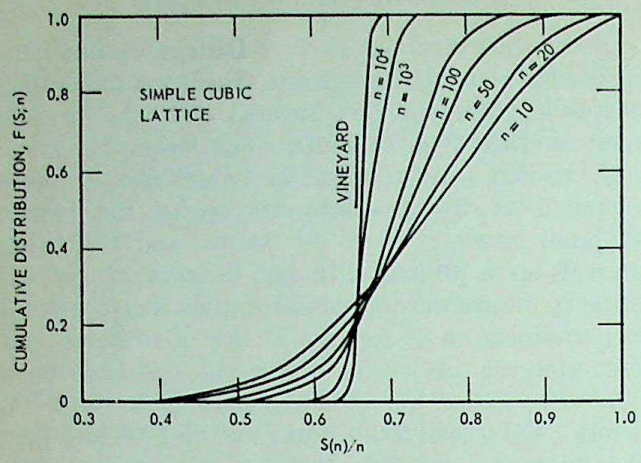


FIG. 2. Distribution function for simple cubic lattice. Vertical tick in Figs. 2–4 indicates Vineyard's asymptotic solution for $S(n)/n$.

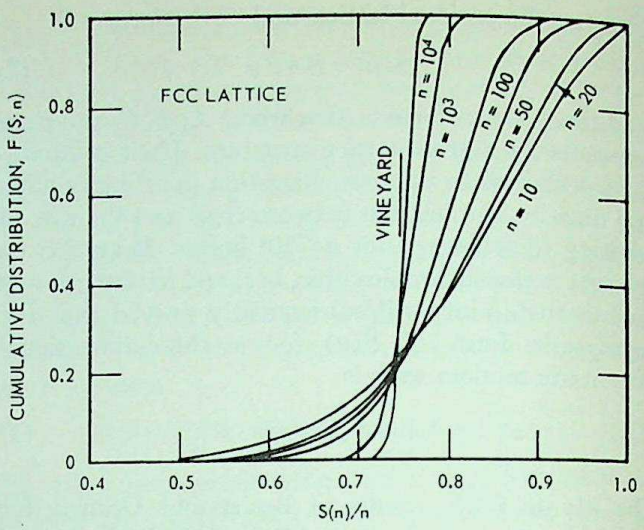


FIG. 4. Distribution function for fcc lattice.

lattices, however, both $S_1(n)/n$ and $S_2(n)/n$ change as n increases, $S_1(n)/n$ tending toward $\bar{S}(n)/n$ from below and $S_2(n)/n$ approaching it from above. The vertical tick defines the lower limit of $\bar{S}(n)/n$ given by Vineyard's calculations for $n\rightarrow\infty$. For all practical purposes, one can consider this asymptotic behavior to be attained at $n=10^4$ jumps. Table I compares the 1000-history Monte Carlo results for $d\bar{S}/dn$ at $n=10^4$ with Vineyard's evaluations for the asymptotic case.

All distributions were skewed negatively for $n\leq 10^4$ as shown by the coefficient of skewness listed in Table II. This skewness is evident from inspection of the curves in Figs. 1–4 for $n\leq 50$. By definition,[6]

$$\gamma_1=\mu_3/\sigma^3, \tag{5}$$

where μ_n is the nth moment of the distribution about the mean. A negative γ_1 means that the tail of the distribution for $S<\bar{S}$ is longer than the forward tail for $S>\bar{S}$. If $\gamma_1>0$ the opposite is true and $\gamma_1=0$ for the normal distribution. A measure of how the shape of the density function $f(S;n)$, in the vicinity of the mode, differs from that for the normal density function is given by the coefficient of excess[6]

$$\gamma_2=\mu_4/\sigma^4 \tag{6}$$

listed in Table III. If $\gamma_2<0$, $f(S;n)$ is more flat in the vicinity of the mode than the normal density function and more peaked if $\gamma_2>0$. In this regard $\gamma_2<0$ for the sp and sc lattices but $\gamma_2>0$ for bcc and fcc lattices. This suggests that γ_1 is independent of the lattice coordination number z but that $\gamma_2<0$ for $z\leq z_0$ where $6\leq z_0<8$.

A χ^2 test was used to judge the over-all fit of a normal distribution to the Monte Carlo data. The averages of the χ^2 values obtained using 21 deg of freedom (23 intervals) are given in Table IV over 500–900 jumps, 1000–5000 jumps, and 5000–10 000 jumps for each lattice. A composite average for the bcc and fcc distributions appears in the last column. According to Evans,[5] for example, given 21 deg of freedom, χ^2 in the range $13\leq\chi^2\leq 30$ indicates that the data at hand is very probably described by the assumed distribution, i.e., the normal distribution in this case. On the other hand, either $\chi^2<10$ or $\chi^2>36$ indicates that it is very unlikely that the data represents a randomly drawn set of values from the assumed distribution. Excepting the sp lattice, χ^2 decreased with increasing n for $n>500$. All χ^2 values were in the plausible fit range $10\leq\chi^2\leq 36$ for $n>1000$. Both the distribution for the sc lattice and that for the bcc lattice exhibited χ^2 values within the

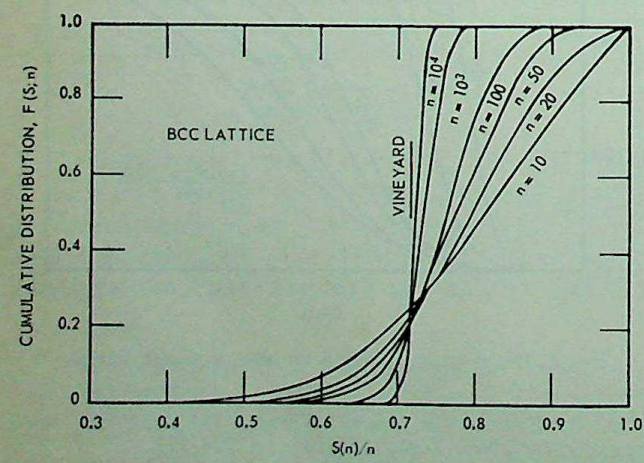


FIG. 3. Distribution function for bcc lattice.

TABLE I. Values of C in $\bar{S}(n)=B+Cn$ given by (a) 1000-history Monte Carlo (MC) calculations at $n=10^4$ and (b) Vineyard's exact solution for $n\rightarrow\infty$.

	(a) Monte Carlo ($n=10^4$)	(b) Vineyard ($n\rightarrow\infty$)	(a)/(b)
sc	0.6641	0.6595	1.005
bcc	0.7216	0.7178	1.005
fcc	0.7472	0.7437	1.005

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Coefficient of skewness (γ_1) for Monte Carlo distributions $F(S;n)$.

n	Skewness SPL	SCL	BCC	FCC
10	−0.054	−0.386	−0.558	−0.663
50	+0.07	−0.286	−0.320	−0.550
100	+0.003	−0.333	−0.364	−0.444
500	−0.130	−0.270	−0.540	−0.300
1000	−0.159	−0.238	−0.418	−0.293
2000	−0.230	−0.350	−0.40	−0.330
3000	−0.440	−0.330	−0.28	−0.320
4000	−0.400	−0.310	−0.26	−0.300
5000	−0.320	−0.320	−0.19	−0.340
6000	−0.320	−0.350	−0.20	−0.340
7000	−0.300	−0.280	−0.19	−0.300
8000	−0.310	−0.230	−0.21	−0.260
9000	−0.370	−0.200	−0.20	−0.220
10 000	−0.382	−0.169	−0.19	−0.205

very probable range for $n \geq 6000$ jumps. Because γ_1 and γ_2 for the bcc and fcc distribution behave similarly above $n=1000$ we feel that a composite average of χ^2 for these two lattices may be meaningful. This composite average is listed in the right-hand column. The χ^2 test indicates: (1) $F(S;n)$ is not well approximated by the normal distribution for $n<1000$; (2) the normal distribution is a good approximation to $F(S;n)$ for $10^3<n<10^4$; (3) $F(S;n)$ is very probably normal for $n>10^4$ in the three-dimensional case.

A least-squares analysis gives the following results for $\bar{S}(n)$ and $\sigma(n)$ on the interval $10^2 \leq n \leq 10^4$:

$$\bar{S}(n)=B+Cn, \tag{7}$$

$$\sigma(n)=An^k. \tag{8}$$

Values for the constants A, B, C, and k appear in Table V.

4. EFFECT OF FLUCTUATIONS ON THE INITIAL ANNEALING RATE

The distributions described in the previous section were computed to supplement a separate analog Monte

TABLE III. Coefficient of excess (γ_2) for Monte Carlo distributions $F(S;n)$.

n	Excess SPL	SCL	BCC	FCC
10	−0.455	−0.355	−0.301	−0.055
50	−0.270	+0.120	−0.14	+0.12
100	−0.158	+0.103	+0.339	+0.216
500	−0.170	+0.210	+0.56	+0.150
1000	−0.048	−0.094	+0.388	+0.043
2000	−0.030	−0.01	+0.76	+0.29
3000	+0.028	−0.10	+0.40	+0.16
4000	+0.22	−0.08	+0.16	+0.21
5000	+0.35	+0.006	+0.13	+0.11
6000	+0.17	−0.10	+0.04	+0.27
7000	−0.05	−0.13	+0.07	+0.34
8000	−0.001	−0.20	+0.11	+0.39
9000	−0.05	−0.20	+0.13	+0.11
10 000	−0.059	−0.12	+0.13	+0.12
	Less slim		More slim	

TABLE IV. Average χ^2 values over Δn ranges indicated. χ_c^2 is the average of columns three and four.

Δn	sp	sc	bcc	fcc	χ_c^2
500–900	27.6	39.8	47.5	38.8	42.0
1000–5000	34.8	34.4	26.8	32.2	31.1
6000–10 000	33.0	22.1	24.7	32.2	26.3

Carlo study on defect annealing in cubic crystallites containing up to 2×10^6 lattice sites. In order to obtain even approximate agreement between the initial annealing rate given by the analog Monte Carlo[7] calculations and that computed using $F(S;n)$, it was necessary to consider the effect of fluctuations in the initial defect concentration within the crystallite. In this section the number of mobile defects remaining after an average of n jumps per defect is computed analytically in two ways. In the first instance the defect encounter probability is computed using $F(S;n)$ and the average concentration over the crystallite; in the second, it is computed using $F(S;n)$ and the average concentrations in each of N_c noninteracting crystallite subvolumes populated according to the random distribution law for M_0 defects in N_c boxes of equal volume. The two expressions are then applied in a numerical example for $M_0=100$ initially mobile defects in a fcc crystallite containing 5×10^5 sites and the results compared with those obtained in the Monte Carlo solution for this system.

A. Average Concentration Approach

Let there exist M_0 mobile defects randomly distributed on Σ lattice sites and M_c defects contained in randomly distributed stable, immobile defect clusters at time $t=0$. Assume that the sole mobile defect removal process is association with other defects at contact. A pair of defects will be taken as stable and immobile in this example. Further, assume that the association probability ϵ at contact is the same when one member of the contact pair is a cluster member as when both defects were mobile just before contact. M_n will denote the number of mobile defects remaining after n jumps per defect. Let $\alpha=1/\Sigma$. The probability

TABLE V. $\bar{S}(n)$ and $\sigma(n)$ given by a least-squares analysis of the Monte Carlo data.

	$\bar{S}(n)=B+Cn$ B	C	$\sigma(n)=An^k$ A	k
sp[a]	...	...	0.194	0.799
sc	16.27	0.6642	0.414	0.608
bcc	16.90	0.7217	0.385	0.604
fcc	15.84	0.7473	0.308	0.635

[a] $\bar{S}(n)=0.7648n^{0.8035}$.

[7] N. R. Baumgardt and J. R. Beeler, Jr., Bull. Am. Phys. Soc. 9, 294 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that a mobile defect will encounter a cluster member on the nth jump is

$$\alpha[M_c+M_0-M_{n-1}]p^*(S;n-1)f(S;n-1), \quad (9)$$

where $p^*(S;n-1)$ is the conditional probability for a defect visiting $(S+1)$ distinct sites on the $(n+1)$th jump provided it had visited exactly S distinct sites in n jumps. The probability that a mobile defect will contact another mobile defect is

$$\alpha[M_{n-1}-1]p^*(S;n-1)f(S;n-1). \quad (10)$$

Contact with a cluster leads to the removal of one mobile defect but contact between two mobile defect leads to a depletion of two mobile defects; hence, the difference equation for M_n is

$$M_n=M_{n-1}-\alpha\epsilon[M_c+M_0-M_{n-1}]\times\sum_{S=S_1(n-1)}^{S=S_2(n-1)}[M_{n-1}f(S;n-1)]p^*(S;n-1)-2\alpha\epsilon[M_{n-1}-1]\times\sum_{S=S_1(n-1)}^{S=S_2(n-1)}[M_{n-1}f(S;n-1)p^*(S;n-1)]. \quad (11)$$

By the definition of $f(S;n)$ the summation reduces to $M_{n-1}\langle p^*(n-1)\rangle$. Equation (11) leads to the differential approximation

$$dM/dn\cong-\alpha\epsilon M\langle p^*(n)\rangle[M+(M_c+M_0-2)]. \quad (12)$$

Integration gives

$$\ln[M/(M+M_c+M_0-2)]_0^n=-\alpha\epsilon(M_c+M_0-2)\int_0^n\langle p^*(n)\rangle dn. \quad (13)$$

Collecting terms, one obtains

$$M(n)=M_0[M_c+M_0-2]/\{[2(M_0-1)+M_c]\times\exp[\alpha\epsilon(M_c+M_0-2)P_n]-M_0\}, \quad (14)$$

where $P_n=\int\langle p^*(n)\rangle dn$. Because $\langle p^*(n)\rangle$ is closely approximated by $d\bar{S}(n)/dn$, we will use

$$P_n=\int_0^n[d\bar{S}/dn]dn=\bar{S}(n). \quad (15)$$

$M(n)$ was evaluated for the particular case $\alpha=2\times10^{-6}$, $M_0=100$, $M_c=0$, and $\epsilon=1$. The results are discussed in Sec. 4C.

B. Consideration of Fluctuations

In the estimation of $M(n)$ outlined above one considered (M_c+M_0) defects in a cell of Σ sites and computed the contact probability from the average concentration

$$\bar{C}=(M_c+M_0)/\Sigma \quad (16)$$

TABLE VI. Probability $g(k)$ to find k defects in a subvolume of 5000 sites when the macroscopic defect concentration is $\bar{C}=2\times10^{-4}$ defects per site. $\bar{C}_k$ is the associated cell average concentration.

k	$g(k)$	$\bar{C}_k\times10^4$
0	0.366032	0
1	0.369730	2
2	0.184865	4
3	0.0609992	6
4	0.0149417	8
5	0.00289778	10
6	0.0046345	12

defects per site. The effect of fluctuations in the initial defect concentration upon the annealing rate can be estimated by introducing the density function $g(k)$ for finding exactly k mobile defects $(k=1,2,\cdots,M_0)$ in a given subvolume of a collection of N_c equal subvolumes with atomic-scale linear dimensions. Let $\tilde{M}(n)$ be the number of mobile defects remaining after n jumps, computed on the basis of $\{g(k)\}$ rather than $\bar{C}$, and set $M_c=0$. $\tilde{M}(n)$ then is

$$\tilde{M}(n)=\sum_{k=1}^{M0}M_k(n). \quad (17)$$

The quantity $M_n(k)$ is obtained by substituting $M_{0k}=kg(k)N_c$ for M_0, and $\alpha_k=1/g(k)\Sigma$ for α, in Eq. (14). The result is

$$M_k(n)=M_{0k}(M_{0k}-2)/\{2(M_{0k}-1)\times\exp[\alpha_k\epsilon(M_{0k}-2)P_n]-M_{0k}\}. \quad (18)$$

$\tilde{M}(n)$ was evaluated for the particular case $\alpha=2\times10^{-6}$, $M_0=N_c=100$ and $\epsilon=1$, using[8]

$$g(k)=M_0!(N_c-1)^{M_0-k}/k!(M_0-k)!N_c^{M_0}. \quad (19)$$

Values of $g(k)$ for $k\leq6$ are given in Table VI along with the associated subvolume concentrations. The average value $\langle k\rangle=1$ corresponds to an average concentration of 2×10^{-4} defects per site over the entire crystallite, the concentration used in computing $M(n)$.

C. Comparison with Monte Carlo Solution

In Fig. 5, $M(n)$, $\tilde{M}(n)$ and Monte Carlo solution, $M_{mc}(n)$, are compared for $\alpha=2\times10^{-6}$, $M_0=N_c=100$, and $\epsilon=1$ for $n\leq1000$. We will assume that M_{mc} is a physically more realistic solution than either M or $\tilde{M}$ because one would expect that it better describes the association of initially closely spaced defects. It will be used as the reference in all discussion. The curves become nearly parallel at $n=500$, and although it is not shown in the figure, they remained nearly parallel up to $n=2000$, the point at which the Monte Carlo runs were terminated. It appears, therefore, that dM/dn is

[8] W. Feller, *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, Inc., New York, 1957), 2nd ed., p. 34.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

a good approximation to the true annealing rate for $n>500$ even though M itself is too large. Rapid association of defects, initially closely spaced, accounts for the more rapid falloff in M_{mc} for $n<500$ relative to that exhibited by M and $\tilde{M}$. The Monte Carlo method enables one to easily describe the removal of these initially close defects during the early annealing stage, a process only crudely described by $\tilde{M}$. It would seem that a rigorous analytical description of fluctuation effects in the association of randomly walking defects on a lattice would be somewhat complicated.

If one accepts M_{mc} as being more realistic than either M or $\tilde{M}$, our results indicate that the use of M in interpreting defect removal data would cause one to overestimate the number of jumps per defect required to attain a given removal fraction.[9] Table VII lists the number of jumps $n(M)$ and $n(M_{mc})$ given by M and M_{mc}, respectively, required to remove a given fraction of mobile defects. Extrapolation of the straight line log-log plot of $n(M)/n(M_{mc})$ versus the removal fraction, gave a limiting value of ~ 1.1 for this ratio. In a strict sense, the ratios listed in Table VII should be regarded as lower bounds, especially when they are associated with $n>1000$. Although each defect was initially contained within a cube of 5×10^5 sites, in the Monte Carlo calculations, some defects migrated out of the cube into a defect-free environment during the annealing process. These escaping defects served to increase Σ a maximum of 2.4% at $n=1000$, i.e., a maximum of 1.2×10^4 sites outside the 5×10^5 site cube were visited. This caused dM_{mc}/dn to be slightly smaller than it would have been had all defects remained inside the 5×10^5 site cell. At $n=2000\Sigma$ was increased a maximum of 3.5% by escaping defects.

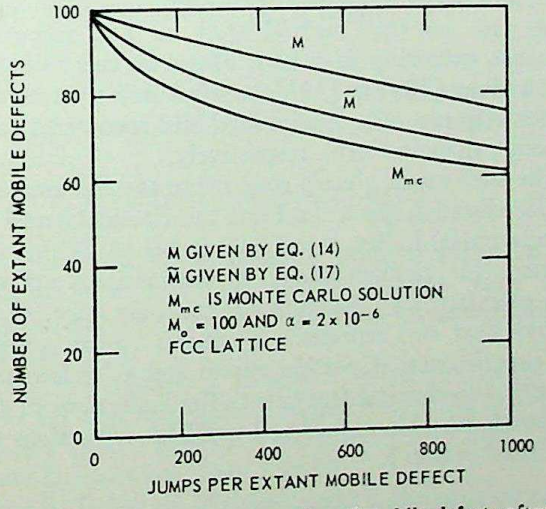


FIG. 5. Comparison of extant number of mobile defects after n jumps as given by M, $\tilde{M}$, and M_{mc} for a fcc lattice.

[9] After completion of this article, experimental evidence supporting this conclusion was found in a paper by F. Dworschak, K. Herschbach, and J. S. Koehler, Phys. Rev. 133, A293 (1964).

TABLE VII. Comparison of the number of jumps $n(M)$ and $n(M_{mc})$ per extant mobile defect, given by M and M_{mc}, respectively, required to remove a given fraction of mobile defects when $M_0=100$ and $\alpha=2\times10^{-6}$.

Fraction removed	$n(\tilde{M})$	$n(M_{mc})$	$n(M)/n(M_{mc})$
0.1	340	60	5.7
0.2	770	200	3.8
0.3	1350	520	7.6
0.4	2300	1070	2.2
0.5	3300	2000	1.6
1.0	...	...	~1.1

Twenty-five Monte Carlo runs of 2000 jumps per extant mobile defect were made to obtain M_{mc}. Five independent runs were made for each of five independent, randomly sampled initial defect distributions in a cube of 5×10^5 sites. A χ^2 analysis of the five initial distributions showed that they were consistent with the density function $g(k)$ of Eq. (19).

5. SUMMARY

The calculations described above indicate:

(1) $F(S;n)$ is not well approximated by the normal distribution for $n<5000$ jumps. However, $F(S;n)$ appears to approach the normal distribution on the interval $5000<n<10\,000$ and to be normal for $n>10^4$ in three-dimensional lattices.

(2) Our results give no conclusive evidence that $F(S;n)$ either does or does not approach the normal distribution in two-dimensional lattices.

(3) The use of Vineyard's asymptotic result for $\bar{S}(n)$, and the assumption that $F(S;n)$ is normal [using Table V for $\sigma(n)$] should give a very good approximation to the distribution of $S(n)$ for $n>10^4$.

(4) A computation of the mobile defect concentration and the *initial* annealing rate must account for the effect of fluctuations in the initial defect concentration.[9] Apparently, however, the annealing rate after the initial annealing stage can be well approximated by considering only the average defect concentration.

ACKNOWLEDGMENTS

The writer is indebted to Dr. G. H. Vineyard for communicating the results of his calculations in advance of publication and thereby suggesting the subject of the present article. J. A. Delaney and Miss N. R. Baumgardt wrote the computer programs used in this work.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Microscopic Dielectric Susceptibility of Germanium

W. COCHRAN
Crystallographic Laboratory, Cambridge, England

AND

J. C. PHILLIPS
Department of Physics, University of Chicago, Chicago, Illinois

(Received 6 January 1964)

The microscopic dielectric susceptibility $\epsilon(q)$ is calculated for germanium using the shell model. It is found that except at $q=0$ the longitudinal and transverse susceptibilities differ and are anisotropic. The results agree qualitatively with Penn's isotropic quantum model based on the random-phase approximation. It is suggested that the quantitative differences may in part represent shortcomings of the RPA in treating the stiffness of covalent bonds.

1. CLASSICAL LATTICE VIBRATION MODELS

MANY classical models have been used to discuss the lattice vibrations of insulators. The interest in such models has been greatly enhanced by the availability of experimentally determined phonon dispersion relations for wavelengths comparable to interatomic separations. Through these models, we may hope to determine certain properties of the valence electron polarizability at short wavelengths.

We may distinguish between two kinds of classical atoms: rigid or polarizable. In each case the valence charge may be regarded as either distributed or localized at a point. For the moment we restrict ourselves to the latter case. The local field corrections in the former case are more involved (Sec. 5).

The rigid point atom yields the simplest models, but is inconsistent with the dielectric properties of ionic crystals. For homopolar Ge, on the other hand, it was natural to suppose that this model would work fairly well. Its failure to fit the dispersion curves determined by inelastic neutron scattering[1] except with forces out to fifth neighbors was shown by Herman.[2]

Polarizable valence electrons can be treated within the adiabatic approximation as charged massless shells. In a free atom the shell and core are connected by an isotropic force constant; in a crystal each shell is also linked to near neighbor cores and shells. This shell model takes account of the fact that an atom can be polarized either through Coulomb or short-range interactions. In germanium displacement of an atom produces a dipole moment on it, but the lowest moment on the atom and its near neighbors is a quadrupole moment. In Lax's theory[3] of the lattice dynamics of germanium, this quadrupole moment is taken to be on the displaced atom; when the shell model is used[4] the interaction involves the component dipoles. This method finds some quantum-mechanical justification in the work of Mashkevich and Tolpygo.[5] It is equivalent to using the atomic displacements and shell displacements (or atomic dipole moments) as generalized coordinates, and writing the energy as a general quadratic form in these coordinates. The shell-shell forces appear formally as the (non-Coulomb) interaction energy of induced dipoles; quantum mechanically they represent the energy required to distort the covalent bonds (Sec. 4).

The microscopic dielectric susceptibility tensor $\epsilon(q)$ is defined by

$$[\epsilon(q)-1]/4\pi=\mathbf{P}(q)/\mathbf{E}(\mathbf{q}). \tag{1.1}$$

Here $\mathbf{E}(\mathbf{q})$ is a macroscopic field of wave number $\mathbf{q}$:

$$\mathbf{E}(\mathbf{q})=\mathbf{E}e^{i\mathbf{q}\cdot\mathbf{r}} \tag{1.2}$$

and $\mathbf{P}$ is the dipole moment per unit cell volume v

$$v\mathbf{P}=\textstyle\sum_i \mathbf{p}_i, \tag{1.3}$$

the sum extending over both atoms in the unit cell. For $\mathbf{q}$ along [100] or [111] axes in cubic crystals, the tensor $\epsilon(\mathbf{q})$ has only longitudinal and transverse components, $\epsilon_l(\mathbf{q})$ and $\epsilon_t(\mathbf{q})$, respectively.

The shell model gives a good fit[4] to the experimental lattice vibration spectrum. From this classical model we calculate $\epsilon_t(q)$ in Sec. 2 and $\epsilon_l(q)$ in Sec. 3. Within the Hartree (or random-phase) approximation, one can also calculate $\epsilon(q)$ quantum mechanically (Sec. 4). It follows that $\epsilon(q)$ represents a natural bridge between the experimental dispersion curves and a fundamental quantum-mechanical treatment. The limitations of this approach due to local field and umklapp corrections are discussed in Sec. 5.

2. TRANSVERSE DIELECTRIC SUSCEPTIBILITY

Within the framework of the shell model, this is a purely classical calculation of the microscopic polarization (1.3) produced by the applied field (1.2). We have[6]

$$Y\mathbf{E}=(\mathbf{S}_t+(Y^2/v)\mathbf{C}_t)\mathbf{W}, \tag{2.1}$$

[1] B. N. Brockhouse and P. K. Iyengar, Phys. Rev. 111, 747 (1958); A. Ghose, H. Palevsky, D. J. Hughes, I. Pelah, and C. M. Eisenhauer, *ibid.* 113, 49 (1959).
[2] F. Herman, Phys. Chem. Solids 8, 405 (1959).
[3] M. Lax, Phys. Rev. Letters 1, 133 (1958); Phys. Chem. Solids 8, 422 (1959).
[4] W. Cochran, Proc. Roy. Soc. (London) A253, 260 (1959).
[5] V. S. Mashkevich and K. B. Tolpygo, Zh. Eksperim. i Teor. Fiz. 32, 520 (1957) [English transl.: Soviet Phys.—JETP 5, 435 (1957)].
[6] W. Cochran, Advan. Phys. 10, 401 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where Y is a constant equal to the charge on a shell. The applied electric field and dipole moments at the two atoms in the unit cell are vector elements of the column matrices $\mathbf{E}$ and $Y\mathbf{W}$. (The elements of $\mathbf{E}$ are therefore equal.) The short-range and Coulomb forces are represented by the 2×2 matrices $\mathbf{S}_t$ and $\mathbf{C}_t$, respectively. Because we are taking $\mathbf{q}$ along a [111] or [100] symmetry direction, and $\mathbf{E}\perp\mathbf{q}$, the matrix elements of $\mathbf{S}_t$ and $\mathbf{C}_t$ are scalar functions of q. For example, for $q\rightarrow0$ all four elements of $\mathbf{C}$ have the value $-\frac{4}{3}\pi$. Different scalar functions enter for $\mathbf{E}\|\mathbf{q}$, so that $\epsilon_t(\mathbf{q})$ calculated here differs from $\epsilon_l(\mathbf{q})$ calculated in the next section.

According to these definitions we can write the polarization (1.3) as

$$v\mathbf{P}=Y\mathbf{W} \tag{2.2}$$

and from (2.1) and (2.2) it follows that

$$\mathbf{P}=((v\mathbf{S}_t/Y^2)+\mathbf{C}_t)^{-1}\mathbf{E}. \tag{2.3}$$

Since for the Ge lattice the off-diagonal matrix elements of $(v\mathbf{S}_t/Y^2+\mathbf{C}_t)$ are complex conjugates, the polarization waves through the two types of atom are not in phase with the applied field $\mathbf{E}$. Because the matrix is Hermitian, one lags as much as the other leads. We may take the sum of the two to define a real microscopic dielectric susceptibility given by

$$\epsilon_{st}(\mathbf{q})=1+4\pi\sum((v\mathbf{S}_t/Y^2)+\mathbf{C}_t)^{-1}, \tag{2.4}$$

where the sum is over all four elements of the matrix.

We can check the correctness of this procedure by evaluating (2.4) at $q=0$. In this limit the shell-shell forces do not affect the polarizability, which depends only on core-shell forces. We find that in terms of the free atom polarizability per unit cell volume β (2.4) reduces to

$$[\epsilon_{st}(0)-1]/[\epsilon_{st}(0)+2]=\tfrac{4}{3}\pi2\beta, \tag{2.5}$$

which is the Clausius-Mossotti formula.[4,7] Thus (2.4) includes local field corrections properly when the valence electrons are idealized as nonoverlapping spherical shells.

Using the values of $\mathbf{C}_t$ and $\mathbf{S}_t$ given previously in Table III and Appendix 4, respectively,[4] one can now evaluate (2.4) along the [100] and [111] axes. The results are shown in Figs. 1 and 2.

A notable feature of these curves is their periodicity, which is not that of the reduced Brillouin zone. The periodicity is a result of the localization of the two valence shells on the atoms rather than merely in the unit cell. We also note that $\epsilon_{st}(q)$ along the [111] direction has a maximum near $\mathbf{q}=K=2\pi a^{-1}$ (111). This may be an umklapp effect (Sec. 5).

[7] The value of β used here is 0.099, which gives $\epsilon(0)=16.0$. It differs slightly from the value of 0.105 used in Ref. 4.

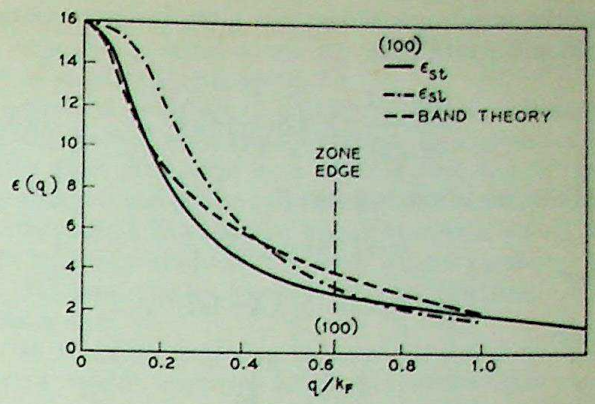


FIG. 1. The longitudinal and transverse shell-model susceptibilities ϵ_{sl} and ϵ_{st}, respectively, as a function of q/k_F along [100] axes, where k_F is the Fermi wave number of a free-electron gas of density equal to the valence electron density in Ge. The curve marked band theory is based on Penn's isotropic band model.

3. LONGITUDINAL SUSCEPTIBILITY

In this case we have

$$Y\mathbf{E}=(\mathbf{S}_l+(Y^2/v)\mathbf{C}_l)\mathbf{W}, \tag{3.1}$$

where as before the elements of the matrix $\mathbf{E}$ are the amplitude of the applied field. The effective field is the sum of the applied, local and macroscopic fields, the two latter being given by $-(Y/v)\mathbf{C}_l\mathbf{W}$. That is, the local and macroscopic fields are not calculated separately by Kellermann's method.[8] At $q\rightarrow0$, for example, the elements of $\mathbf{C}_l$ are all $8\pi/3$. The amplitude of the macroscopic field is -4π times the amplitude of the polarization wave, and this is true for general values of q.[9]

Let P be the amplitude of the polarization wave, E that of the applied field and $E_M=-4\pi P$ that of the macroscopic field. We define $\epsilon_l(q)$ by

$$[\epsilon_l(q)-1]/4\pi=P/(E+E_M),$$

so that

$$[\epsilon_l(q)-1]/4\pi\epsilon_l(q)=P/E. \tag{3.2}$$

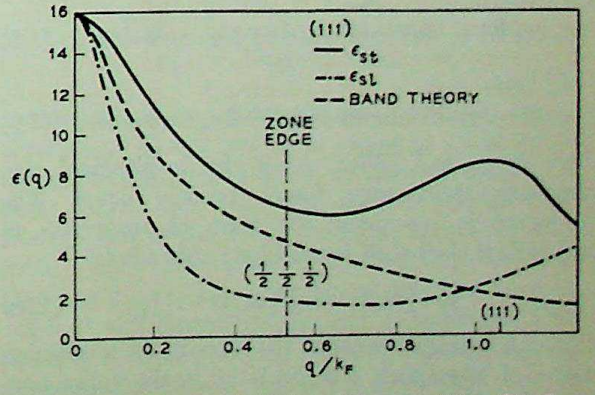


FIG. 2. The same as Fig. 1, but with $\mathbf{q}$ parallel to [111].

[8] E. W. Kellerman, Phil. Trans. Roy. Soc. (London) A238, 513 (1940).

[9] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

By the same argument as was given in the previous section, we then have

$$\frac{\epsilon_l(q)-1}{4\pi\epsilon_l(q)}=\sum\left(\mathbf{S}_l\frac{v}{V^2}+\mathbf{C}_l\right)^{-1}. \tag{3.3}$$

Making use of the fact that if

$$(\epsilon-1)/4\pi=\sum \mathbf{X}^{-1},$$

then

$$(\epsilon-1)/4\pi\epsilon=\sum(\mathbf{X}-4\pi\mathbf{I})^{-1},$$

where

$$\mathbf{I}=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}.$$

We finally have

$$\epsilon_l(q)=1+4\pi\sum(\mathbf{S}_l(v/V^2)+\mathbf{C}_l-4\pi\mathbf{I})^{-1}. \tag{3.4}$$

The elements of $\mathbf{C}_l$ and of $\mathbf{C}_t$ are related by[8]

$$\mathbf{C}_l+2\mathbf{C}_t=0. \tag{3.5}$$

It is readily verified from (2.4), (3.4), and (3.5), together with the fact that $S_t(0)=S_l(0)$, that $\epsilon_l(0)=\epsilon_t(0)$ (as required by cubic symmetry).

Using the expressions for $\mathbf{S}$ appropriate to longitudinal modes,[4] we have evaluated (3.5) along the [100] and [111] symmetry directions. The results are shown in Figs. 1 and 2, where they are compared with $\epsilon_{st}(q)$. Except at $q=0$ (where shell-shell forces play no role), $\epsilon_{sl}(\mathbf{q})$ is not equal to $\epsilon_{st}(\mathbf{q})$; the difference is small along the [100] direction where $\mathbf{q}$ makes equal angles with all shell-shell force constants, and large along [111] directions where $\mathbf{q}$ is parallel to one of the shell-shell force constants.

The idea that shell-shell forces are chiefly responsible for the difference between $\epsilon_{sl}(\mathbf{q})$ and $\epsilon_{st}(\mathbf{q})$ is borne out by the explicit expressions[4] for $-S/S_0$. These are

[100] trans. $\cos qa/4-i\gamma_s\sin qa/4$, (3.6)

[100] long. $\cos qa/4$, (3.7)

[111] trans.

$$\cos^3\theta+\gamma_s\sin^2\theta\cos\theta+i(\sin^3\theta+\gamma_s\cos^2\theta\sin\theta), \tag{3.8}$$

[111] long.

$$\cos^3\theta-2\gamma_s\sin^2\theta\cos\theta+i(\sin^3\theta-2\gamma_s\cos^2\theta\sin\theta), \tag{3.9}$$

where $\theta=qa/4\sqrt{3}$. With $\gamma_s=0$ the longitudinal and transverse short-range forces become equal. The parameter γ_s measures[4] the ratio of transverse to longitudinal shell-shell forces along cube axes:

$$\gamma_s=[\phi_{xy}(s-s)]/[\phi_{xx}(s-s)], \tag{3.10}$$

where ϕ_{xy} is proportional to the force in the x direction on shell 1 when shell 2 is displaced in the y direction. The value of γ_s used previously[4] was 0.7.

One can express γ_s in terms of the ratio of bond bending to bond stretching forces. One obtains

$$\gamma_s=(k_s-k_B)/(k_s+2k_B). \tag{3.11}$$

Thus, $\gamma_s=0.7$ corresponds to $k_B/k_s=0.13$. This is in quantitative agreement with the ratio of core-core and core-shell bond bending and bond stretching forces required to fit the elastic constants[4] and other theoretical[11] and experimental data[12] in the $q=0$ limit.

4. QUANTUM-MECHANICAL SUSCEPTIBILITY

In treating the real crystal quantum mechanically, we are concerned with fields throughout the unit cell and not merely at the atomic sites. If the external potential is $v(x)$ and the self-consistent total potential is $V(x')$, we have the general relation

$$v(x)=\frac{1}{\Omega}\int\mathfrak{D}(x,x')V(x')dx'. \tag{4.1}$$

Here the integral extends over the volume of the crystal.

The microscopic dielectric susceptibility is $\mathfrak{D}(x,x')$. When (4.1) is Fourier analyzed, taking advantage of periodicity, one obtains[13–15]

$$v(\mathbf{q},\mathbf{K}_p)=\sum_{p'}\mathfrak{D}(\mathbf{q},\mathbf{K}_p,\mathbf{K}_{p'})V(\mathbf{q},\mathbf{K}_{p'}). \tag{4.2}$$

We are interested in the case $\mathbf{K}_p=0$. Then (4.2) may be rewritten

$$v(q,0)=\epsilon(q)V(q)+\sum_{p'}\mathfrak{D}(\mathbf{q},0,K_{p'})V(\mathbf{q},\mathbf{K}_{p'}), \tag{4.3}$$

where the prine on the summation excludes $\mathbf{K}_{p'}=0$.

We have chosen to regard $\epsilon(q)$, the $\mathbf{K}_p=\mathbf{K}_{p'}=0$ component of the operator $\mathfrak{D}$, as the quantum-mechanical analog of $\epsilon_s(q)$. Because of the point nature of the shell model, this choice is not unique. We shall see in Sec. 5 that with this choice the local field corrections to $\epsilon(q)$ associated with a distributed shell charge are small, at least within the first Brillouin zone. We also postpone to Sec. 5 the discussion of the second set of terms on the right of (4.3), which may be called umklapp corrections.

If one knows one-electron wave functions and energies throughout the Brillouin zone, $\epsilon(q)$ may be calculated using the random-phase approximation[16] (RPA). For the static dielectric function we are considering,

$$\epsilon(\mathbf{q})=1+\frac{1}{N}\sum_n\frac{f_{0n}(\mathbf{q})\omega_p^2}{\omega_{n0}^2}. \tag{4.4}$$

The oscillator strengths $f_{0n}(\mathbf{q})$ are, in general, different for longitudinal and transverse polarizations; they are

[10] A. D. B. Woods, W. Cochran, and B. N. Brockhouse, Phys. Rev. 119, 980 (1960).
[11] L. Kleinman, Phys. Rev. 128, 2614 (1962).
[12] A. Segmuller, Phys. Letters 4, 277 (1963).
[13] S. L. Adler, Phys. Rev. 126, 413 (1962).
[14] N. Wiser, Phys. Rev. 129, 62 (1963).
[15] M. Azuma, J. Phys. Soc. Japan 18, 194 (1963).
[16] P. Nozieres and D. Pines, Phys. Rev. 109, 762 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the same only in the case of an isotropic solid such as a free-electron gas.

An isotropic model of a semiconductor has been studied by Penn.[17] The model contains an energy gap between the valence and conduction bands. Near this gap the Bloch functions are made out of two plane waves, $\mathbf{k}$ and $\mathbf{k}-\mathbf{K}$. The "reciprocal lattice vectors" $\mathbf{K}$ involved are always parallel to $\mathbf{k}$. The umklapp terms are negligible, being of order N^{-1}, where N is the number of atoms in the crystal. The model also makes the longitudinal and transverse dielectric susceptibilities equal. This result may apply more generally than merely to Penn's model. That is, the anisotropy of $\epsilon(\mathbf{q})$ may be a measure of the magnitude of umklapp corrections.

Penn shows that a dielectric susceptibility consistent with his band model is given by the formula

$$\epsilon(q)=1+\left(\frac{\hbar\omega_p}{E_g}\right)^2\frac{1}{[1+(E_F/E_g)(q/k_F)^2]^2}. \quad (4.5)$$

Here E_F and k_F are the Fermi energy and wave number of the valence electrons regarded as a free-electron gas. With $E_F=12$ eV, $E_g=4$ eV, (4.5) reduces to

$$\epsilon(q)=1+\frac{15}{[1+3.0(q/k_F)^2]^2}. \quad (4.6)$$

We have called (4.6) the dielectric function obtained from band theory. It is also plotted in Figs. 1 and 2.

We see that along the [100] axis, where the difference between ϵ_{st} and ϵ_{sl} is small, $\epsilon(q)$ from (4.6) agrees well with ϵ_s. This confirms our identification of $\epsilon(\mathbf{q})$ with the $\mathbf{K}_p=0$ component of $\mathfrak{D}(\mathbf{q},\mathbf{K}_p)$. That the agreement of a classical model with point charges with a quantum-mechanical band model should be so good is at first unexpected. We are less surprised, however, when we remember that the shell model gives a good fit to the neutron data in a rather simple way.

Along the [111] axis, ϵ_{st} differs substantially from ϵ_{sl}, even within the zone. (At the zone edge, $\epsilon_{st}\simeq 3\epsilon_{sl}$). Within the zone, ϵ falls between ϵ_{sl} and ϵ_{st}. We now wish to consider whether the large difference between ϵ_{sl} and ϵ_{st} is real, and if so, what shortcoming of the band model makes the difference zero there.

To answer this question we must discuss the quantum-mechanical equivalent of the parameter γ_s. According to (3.11), $S_0\gamma_s$ increases with k_{ss}, the force constant involved in stretching the covalent bond. The tendency of shell-core forces is to pull the valence charge distributions apart; it is necessary to balance this with a bond stretching force. Quantum mechanically, bond stretching causes a substantial lowering of valence charge density near the covalent plane, defined as the plane midway between the two atoms in the unit cell. This requires an increase in kinetic energy, which is represented by the shell-shell force constant. The charge density at the covalent plane is little altered by bond bending so that $k_b \ll k_s$.

We can now see that the quantum-mechanical generalization of the bond stretching force constant k_s is a force distributed over the covalent face of the atomic cell and acting on a distributed and deformable valence shell. The effect of these refinements should be to reduce the differences between ϵ_{st} and ϵ_{sl} along the [111] axes, perhaps by 50%, but not eliminate them altogether.

It would be interesting to know whether a realistic band model would produce a substantial difference between ϵ_{st} and ϵ_{sl}. From previous experience[18] with $\epsilon(q)$ at $q=0$, it appears that much ingenuity would be required to carry out calculations for $q\neq 0$. It also appears that with many bands, the tendency towards isotropy within the random-phase approximation (4.3) is strong. We suspect that the anisotropy manifested in Fig. 2 indicates that covalent bonding involves short-range correlations which are not treated accurately by the RPA.

5. LOCAL FIELD AND UMKLAPP CORRECTIONS

It was remarked in Sec. 2 that $\epsilon_s(q)$ is periodic along [100] and [111] axes. If we replace a point shell at $\mathbf{R}$ by a specified charge distribution $\sigma(\mathbf{r}-\mathbf{R})$, we find that the force on a shell is

$$\mathbf{E}\int e^{i[\mathbf{q}\cdot(\mathbf{R}+\mathbf{r})]}\sigma(\mathbf{r})d^3\mathbf{r}. \quad (5.1)$$

This is just the previous value, multiplied by the form factor of $\sigma(\mathbf{r})$, which we write as

$$Y(\mathbf{q})=\int\sigma(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}d^3\mathbf{r}. \quad (5.2)$$

On Fourier analyzing the polarization one finds, e.g., for ϵ_{st},

$$\frac{\epsilon_{st}(q)-1}{4\pi}=\frac{Y^2(q)}{Y^2(0)}\sum\left(\frac{v\mathbf{S}_t}{Y^2}+\mathbf{C}_t\right)^{-1}. \quad (5.3)$$

When an extended charge distribution is used the elements of $\mathbf{C}_t$ will not be the same as those we have met previously, which apply to charges localized at a point (or to nonoverlapping spherical distributions). For example, as the charge is spread out, $C_t(0)$ changes from $-4\pi/3$ to zero, and $C_l(0)$ from $+8\pi/3$ to $+4\pi$. However these changes can be taken up in $\mathbf{S}$, in other words it will always be possible to keep $(v\mathbf{S}/Y^2)+\mathbf{C}$ unchanged. Thus with this quantity unchanged, so that the phonon dispersion curves are unchanged, the right-hand side of (5.3) is multiplied by the damping factor $Y^2(q)/Y^2(0)$ compared to (2.4).

[17] D. Penn, Phys. Rev. 128, 2093 (1962).

[18] D. Brust, J. C. Phillips, and F. Bassani, Phys. Rev. Letters 9, 89 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

By comparing $\epsilon_s(q)$ with $\epsilon(q)$, we can see that the damping factor must be nearly 1 within the Brillouin zone and start to decrease rapidly for $q \gtrsim k_F$. Outside the zone $\epsilon_s(q)$ represents an analytic continuation of the susceptibility which gives a good fit to that deduced from the phonon dispersion curves inside the zone. The continuation is not unique, as we see from the presence of the damping factor in (5.3). Nevertheless, the damping is probably not so rapid as to destroy the maxima in ϵ_{st} and ϵ_{sl} for $\mathbf{q}$ near $\mathbf{L}=[111]$ reciprocal lattice vector.

Such maxima can be called umklapp effects. Azuma[15] has shown that the umklapp terms in (4.2) lead to an effect of this sort. He finds

$$\epsilon(\mathbf{K}_p)=\mathfrak{D}(\mathbf{K}_p,0)/[1-C_p/\epsilon(0)], \qquad (5.4)$$

where C_p is a number of order unity. The enhancement predicted by (5.4) may be related to the subsidiary maxima found in ϵ_s for $\mathbf{q}$ near $\mathbf{L}$.

6. CONCLUSIONS

In spite of the differences between a classical model of point valence shells and a quantum-mechanical model with valence charge spread over the atomic cell, the microscopic dielectric susceptibility obtained from the shell model is in qualitative agreement with the RPA result, especially within the zone. This provides some microscopic justification of the shell model.

Lax[19] has criticized the shell model on the grounds that its point nature grossly overemphasizes the local field corrections. He cites Stark shifts of nuclear resonances,[20] which place an upper limit on the corrections of order 50%, whereas the point charge model implies corrections by a factor $(\epsilon+2)/3\sim 5$, i.e., ten times larger than the experimental shifts.

Our calculations here demonstrate how the shell model overcomes these difficulties. The microscopic susceptibility calculated from band theory, $\epsilon(q)$, implies small (or, in the case of the isotropic model, no) local field corrections. Nevertheless $\epsilon_s(q)$ agrees well with $\epsilon(q)$ near the zone edge. This is because the short-range forces dominate both the lattice vibration frequencies and $\epsilon_s(q)$ in this region. As $q\to 0$ phase differences between nearest neighbors are reduced, the short-range forces are reduced, and the long-range Coulomb forces become important. In this limit the nature of the charge distribution in the unit cell is irrelevant, so long as it is treated *consistently* to give the right limit,

$$\lim_{q\to 0}\epsilon_{st}(q)=\lim_{q\to 0}\epsilon_s(q)=\epsilon_0=16, \qquad (6.1)$$

for the dielectric properties. This the shell model does; adjustment of β to give $\epsilon_0=16$ automatically renormalizes the theory.

We can now reverse this argument and use the fact that short-range forces dominate near the zone edge to argue that the differences between $\epsilon_{st}(q)$ and $\epsilon_{sl}(q)$ along [111] axes are real. These differences may represent the first observable evidence for departures from the RPA in metals or semiconductors.

ACKNOWLEDGMENTS

One of us (J. C. P.) has benefited from conversations with M. Lax. Part of this work was carried out by J. C. P. as a Guggenheim fellow with a Sloan grant in aid at Cambridge, and part at Bell Laboratories.

[19] M. Lax, 1963 Copenhagen Conference on Lattice Dynamics (to be published).

[20] N. Bloembergen, *Proceedings of the Eindhoven Conference on Magnetic and Electric Resonance and Relaxation* (North-Holland Publishing Company, Amsterdam, 1963), p. 39.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 134, NO. 6A 15 JUNE 1964

Bulk Superconductivity in Dilute Hexagonal Titanium Alloys

F. HEINIGER* AND J. MULLER*
Department of Solid State Physics, Federal Institute of Technology, Zurich, Switzerland
(Received 31 January 1964; revised manuscript received 2 March 1964)

The superconductivity, the low-temperature specific heat, and the paramagnetic susceptibility have been investigated in a series of dilute hexagonal titanium alloys. In the solid solutions with V or Nb, both the transition temperature and the density of states are strongly increased and the entire volume exhibits superconductivity. A similar study of Ti-Fe alloys suggests either an incomplete volume transition or a different entropy variation in the superconducting state.

INTRODUCTION

THE significance of the enhancement of the superconducting critical temperature, which is observed in dilute alloys of titanium with other transition elements, has been subject to considerable controversy.[1–7] Matthias *et al.* interpreted the behavior of these alloys as evidence of the existence of a mechanism of a different type from that provided by the electron-phonon interaction,[1,8] while recent work by Cape[6] on Ti-Mn alloys suggested that the high transition temperatures might simply be due to filamentary cubic precipitations in the hexagonal matrix. The experiments to be described indicate (a) that in dilute solid solutions of Ti with V or Nb, in which cases the hexagonal phase extends to the highest solute concentrations, the enhancement of superconductivity is a real bulk effect; and (b) that the substantial increase of the transition temperature is accompanied by a marked increase of both the electronic specific heat and the magnetic susceptibility. On the other hand, dilute alloys of Ti with Fe show a different calorimetric behavior at the temperature of the diamagnetic transition, and the normal electronic specific heat remains essentially unchanged.

HEXAGONAL ALLOYS OF Ti WITH V AND Nb

The structure of Ti-rich alloys is goverened by the $\alpha-\beta$ phase transition and the Martensite-type diffusionless transformation. The hexagonal alloys containing V or Nb were obtained by quenching after repeated melting in the argon arc. According to x-ray examination, the quenched specimens were single phase up to more than 10 at.% in both cases. The initial slope of the superconducting transition temperature T_c versus solute concentration is practically identical in the V and Nb series, the increase near pure Ti being approximately 1°K per at.%. The observation that very rapid quenching yields somewhat higher transition temperatures has to be interpreted in terms of the state of internal strains or a high concentration of defects, since we have found that Ti-Zr alloys show a similar effect.

Bearing in mind that even a complete diamagnetic transition may not unambiguously indicate a volume effect, we have to rely on the calorimetric transition which is insensitive to possible partial screening. The specific heat of representative samples weighing 3–10 g (small specimens are essential for efficient quenching[9])

* Present address: Institut de Physique Expérimentale, Université de Genève, Geneva, Switzerland.

[1] B. T. Matthias, V. B. Compton, H. Suhl, and E. Corenzwit, Phys. Rev. 115, 1597 (1959).

[2] E. Bucher and J. Muller, Helv. Phys. Acta 34, 410 (1961).

[3] W. Buckel, G. Dummer, and W. Gey, Z. Angew. Phys. 14, 703 (1962).

[4] C. J. Raub and C. A. Andersen, Z. Physik 175, 105 (1963).

[5] J. A. Cape and R. R. Hake, Bull. Am. Phys. Soc. 8, 192 (1963); R. R. Hake and J. A. Cape, *ibid.* 8, 419 (1963).

[6] J. A. Cape, Phys. Rev. 132, 1486 (1963).

[7] R. L. Falge, Phys. Rev. Letters 11, 248 (1963).

[8] See, for example, B. T. Matthias, in *Proceedings of the Eighth International Conference on Low Temperature Physics, London, 1962* (Butterworths Scientific Publications Ltd., London, 1963), p. 135.

[9] Our specimens are quenched from slightly below the solidus temperature in an intense stream of cold argon. In another case where the quenching rate is critical [E. Bucher, F. Laves, J. Muller, and H. von Philipsborn, Phys. Letters 8, 27 (1964)], this method has proved to yield better results than water quenching.

Copyright © 1964 by The American Physical Society.
CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

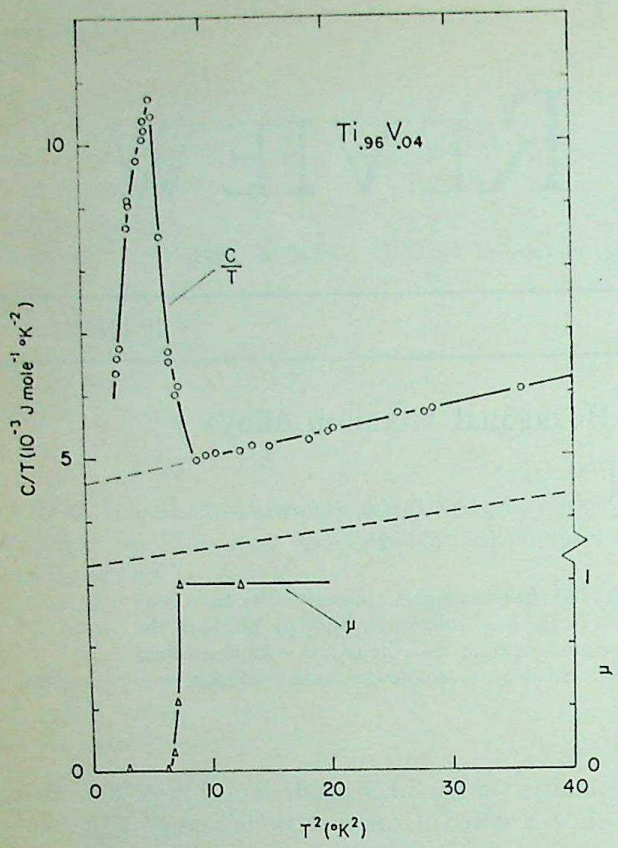


FIG. 1. Specific heat (C/T versus T^2) and magnetic permeability μ of hexagonal $Ti_{0.96}V_{0.04}$. Lower broken line indicates C/T versus T^2 of pure Ti. For T_c versus concentration, see Ref. 2.

have been determined with an absolute accuracy of $\pm 2\%$.[10] Figures 1 and 2 show the results for two alloys (single-phase hcp) containing 4 at.% V or Nb, respectively. Considering the discontinuity at the superconducting transition $[(C_s-C_n)/\gamma T]_{T=T_c}$, we conclude that the bulk of the material is superconducting below the temperature of complete diamagnetic shielding. Therefore, the marked increase of the magnetically observed transition temperature,[2] compared to T_c of pure Ti, is characteristic of the solid solution and is not due to precipitations. Furthermore, as it is seen from Table I, the V and Nb solutes significantly affect the normal electronic specific heat. We also note that the change of γ is much more pronounced than the change of χ. It is reasonable to assume the presence of a large paramagnetic contribution in addition to the spin susceptibility, although in regard to different interaction corrections, γ and χ_{spin} alone need not vary similarly upon alloying. A plot of $\ln(T_c/\theta)$ versus $1/\gamma$ reveals that the data for dilute Ti-V and Ti-Nb essentially agree with those obtained with the isoelectronic Ti-Zr system.[11] Such a relationship is expected qualitatively on the basis of the BCS theory. For the time being, a more detailed interpretation is of little use since the relative importance of the effective Coulomb interaction is not known and no information on the isotope effect in group IV transition elements is available.

[10] Jean Muller, Helv. Phys. Acta 33, 514 (1960).

[11] E. Bucher, F. Heiniger, J. Muheim, and J. Muller, Rev. Mod. Phys. 36, 146 (1964).

TABLE I. Variation of electronic specific heat γ, Debye temperature θ, and magnetic susceptibility χ, of hexagonal alloys of Ti with V and Nb.

	Ti	$Ti_{0.96}V_{0.04}$	$Ti_{0.96}Nb_{0.04}$
γ (10^{-3} J mole^{-1} °K^{-2})	3.3	4.6	4.3
θ_0 (°K)	415	365	340
χ_{at}/χ_{at} (Ti)	(1)	1.09	1.05
χ_{at} (10°K)/χ_{at} (300°K)	0.92	0.92	0.92

In fact, a knowledge of the isotope effect from experiment would permit separation of the interaction into a Coulomb and a phonon part.[12] Unfortunately, attempts of determining the isotopic shift of the transition temperature in[13] Ti and[14] Zr have been unsuccessful due to the difficulties of the sample preparation.

DILUTE Ti-Fe ALLOYS

For comparison, several dilute Ti-Fe alloys with iron content up to 1.5 at.% have been prepared in the same

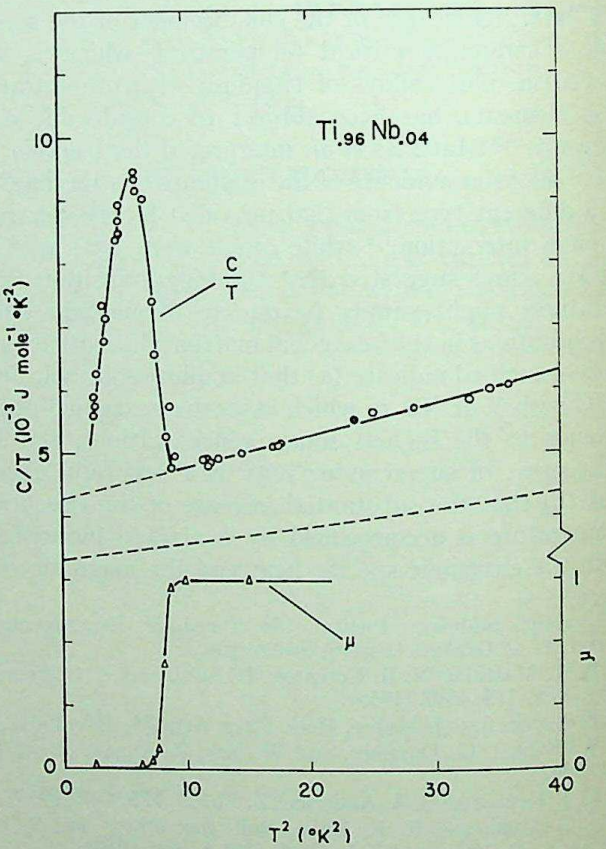


FIG. 2. Specific heat and permeability of hexagonal $Ti_{0.96}Nb_{0.04}$.

[12] E. Bucher, F. Heiniger, and J. Muller, Phys. Kondens. Materie 2, 210 (1964).

[13] R. G. Netzel and J. R. Dillinger, *Proceedings of the 7th International Conference on Low Temperature Physics and Chemistry, Toronto, 1960* (University of Toronto Press, Toronto, 1961), p. 389.

[14] E. Bucher, J. Muller, and J. L. Olsen (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

way as the Ti-V and Ti-Nb specimens. X-ray diffraction showed only hexagonal lines in the quenched state. Optical micrographs closely resembled the one obtained by Cape[6] for as-cast $Ti_{0.99}Mn_{0.01}$. No evidence of a pronounced precipitation of the β phase could be detected. From 0.6 to 1.5 at.% Fe, the midpoint temperature of the change in permeability, observed ballistically in a field of 1 Oe, was found to increase linearly with concentration, in agreement with the earlier data.[1] However, in contrast to the complete diamagnetic transition, the specific heat shows only a slight anomaly (Fig. 3). Similar discrepancies between the amount of flux exclusion and the calorimetric transition have been observed in a number of cases as, e.g., in the σ phase of the Nb-Ir system.[12] A smaller than expected calorimetric anomaly has also been noted by Morin[15] at the apparent T_c of $Ti_{0.99}Fe_{0.01}$. According to the arguments given below, the hypothesis that the observed effect is caused by minor body-centered cubic inclusions seems very improbable, although we cannot rule out completely that traces of the latter may be present. Rather, we are left with two possible interpretations assuming that the transitions do occur in the hexagonal phases (α and α'). Unlike the behavior of Mn in α-Ti,[5-7] the present specific heat data and susceptibility measurements[6] indicate no localized moments on the iron solute.

If we assume a normalized distribution function $f(T_c)$ of the actual volume fraction undergoing a transition at T_c, the observed electronic specific heat will be given by

$$C(T)=\int_T^{T_{c\max}} C_s(T,T_c)f(T_c)dT_c+\int_0^T \gamma(T_c)Tf(T_c)dT_c,$$

where

$$\int_0^{T_{c\max}} f(T_c)dT_c=1.$$

For simplicity, we define

$$x=\int_{T_1}^{T_2} f(T_c)dT_c,$$

with T_2 representing the upper onset of the transition, and $T_1=T_2/2$. Further, we insert the value for $C_s(T,T_c)$ provided by the two-fluid model

$$C_s(T,T_c)=[3\gamma(T_c)/T_c^2]T^3,$$

[15] F. J. Morin (unpublished), quoted in Ref. 6.

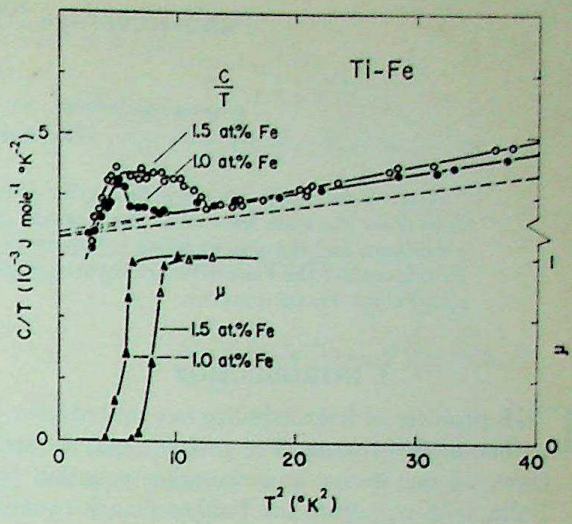


FIG. 3. Specific heat of dilute Ti-Fe alloys. Note complete magnetic transition (μ).

since in the region between T_c and $T_c/2$ this is, for many superconductors, a better approximation than the corresponding BCS function. By the use of a numerical trial and error analysis, we obtain reliably $0.95<x<1$ for the V and Nb alloys, but $0.3<x<0.35$ for Ti+1.5 at.% Fe. If, in fact, only 30% of the volume were superconducting in the latter, a core structure with appreciable microscopic concentration gradients would be required to explain the magnetic transition. On the other hand, this figure almost excludes the possibility of a transition just due to cubic precipitations which are not detected in the x-ray patterns, unless induced superconductivity in the environment of such precipitations were important.

An alternative interpretation of the reduced calorimetric anomaly would have to be based on the assumption that the entropy of the superconducting state is lowered much less than usual, with a corresponding different temperature variation of the energy gap. But if so, it would be difficult to understand why isostructural and isoelectronic dilute Ti-V and Ti-Fe solid solutions, both without localized magnetic states, should exhibit a different kind of superconducting transition.

ACKNOWLEDGMENTS

We wish to thank Dr. E. Bucher, D. Bender, A. Menth, and H. von Philipsborn for their assistance. This work has been supported by the Eidg. Kommission zur Förderung der Forschung and the Schweiz. Nationalfonds.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Time-Reversed Motion in Kinetic Theory*

M. B. Lewis
Technological Institute, Northwestern University, Evanston, Illinois

(Received 22 January 1964)

A set of reversible equations for F_1, the first distribution function and g, the correlation function, are derived for the weak force case. The "forward motion," i.e., development in time from uncorrelated initial conditions, and the corresponding reverse motion are examined. In the "forward motion" the equation for F_1 evolves into the Fokker-Planck equation while in the corresponding reverse motion F_1 is described by an anti-Fokker-Planck equation.

I. INTRODUCTION

THE problem of irreversibility has existed since the time of Boltzmann. The problem may be stated as: How can one derive an irreversible equation (e.g., the Boltzmann equation, the Fokker-Planck equation) on the basis of reversible mechanics? In recent years a number of derivations of irreversible equations have been accomplished based on a variety of assumptions. However, these methods have not completely illuminated the transition from the reversible to the irreversible equations. Most of these methods start from the Bogoliubov-Born-Green-Kirkwood-Yvon (BBKGY) equations,[1] which for an infinite system are

$$\frac{\partial F_s(x_1,\cdots x_s,t)}{\partial t}+\mathfrak{IC}_s F_s(x_1,\cdots x_s,t)$$
$$=-\frac{1}{v}\int dx_{s+1}\sum_{i\leq s}\theta_{i,s+1}F_{s+1}(x_1,\cdots x_{s+1},t) \quad (1)$$
$$s=1,2,\cdots,$$

where $x_i\equiv\{\mathbf{q}_i,\mathbf{p}_i\}$; $\mathbf{q}_i$, $\mathbf{p}_i$ being the position and momentum of the ith particle,

$$\theta_{ij}\equiv\frac{\partial\varphi_{ij}}{\partial\mathbf{q}_i}\cdot\frac{\partial}{\partial\mathbf{p}_i}+\frac{\partial\varphi_{ij}}{\partial\mathbf{q}_j}\cdot\frac{\partial}{\partial\mathbf{p}_j}$$

and $\varphi_{ij}(|\mathbf{q}_i-\mathbf{q}_j|)$ is the intermolecular potential. $\mathfrak{IC}_s$ is given by

$$\mathfrak{IC}_s(x_1,\cdots x_s)\equiv\sum_{i=1}^{s}\frac{\mathbf{p}_i}{m}\cdot\frac{\partial}{\partial\mathbf{q}_i}-\sum_{i<j}^{s}\theta_{ij},$$

and v is the volume per particle. F_s is defined through $D_N(x_1,\cdots x_N,t)$ the probability distribution in Γ space for the entire system by

$$F_s(x_1,\cdots x_s,t)$$
$$=V^s\int\cdots\int D_N(x_1,\cdots x_N,t)dx_{s+1}\cdots dx_N. \quad (2)$$

The sth equation of (1) is obtained by integrating the Liouville equation

$$(\partial D_N/\partial t)+\mathfrak{IC}_N D_N=0 \quad (3)$$

over the coordinates $x_{s+1},\cdots,x_n$.

In particular, the method of Bogoliubov[2] is to assume that all $F_s(s\geqslant 2)$ are functionals of F_1, and the form of the functionals are determined by an assumed boundary condition on the functionals. On this basis, the first BBKGY equation [Eq. (1), $s=1$] becomes the Boltzmann equation (or the Fokker-Planck equation depending on whether the expansion parameter for the distribution function is v^{-1} or the strength of the interaction). The assumptions of Bogoliubov obscure the transition from the set of reversible equations (1) to the irreversible equations. In fact Cohen and Berlin[3] have shown that, on the basis of an equally plausible boundary condition on the functionals, the collision term will be the negative of the Boltzmann collision term (an anti-Boltzmann equation). On the other hand, Green and Piccirelli[4] have shown that, on the basis of a product type condition on $F_s(s\geqslant 2)$ at *the initial time*, the higher distribution functions do, in the course of time, become the functionals of F_1 predicted by Bogoliubov. From this point of view it is no longer important that the Boltzmann equation is time irreversible since the Boltzmann equation evolves from the first BBKGY equation only after some period of time providing the initial $F_s(s\geqslant 2)$ fall within the assumed class of initial conditions. The reverse motion is presumably accomplished, since the BBKGY equations are reversible, by another special class of initial $F_s(s\geqslant 2)$; this motion is probably describable by a single equation for F_1 (such as the anti-Boltzmann equation).[5,6] It is to these points that this paper is devoted.

* Supported by the U.S. Office of Naval Research.

[1] G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, Rhode Island, 1963).

[2] N. N. Bogoliubov, *Problems of a Dynamical Theory in Statistical Physics*, translation by E. K. Gora, *Studies in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962).

[3] E. G. D. Cohen and T. H. Berlin, Physica **26**, 717 (1960).

[4] M. S. Green and R. A. Piccirelli, Phys. Rev. **132**, 1388 (1963).

[5] There remains the question of how one can justify the use of the Boltzmann equation when other types of developments are possible with other types of initial conditions. The answer must be in the following: If we consider a system in which all that is known initially is F_1, then there exists a large class of initial D_N's that give the same F_1 but different $F_s(s\geqslant 2)$. In the spirit of statistical mechanics we can assign a probability to these D_N's and obtain a probability distribution for F_s and ask for the most

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We consider a special case, namely a weak potential in a spatially uniform system. This case is general enough to produce an "irreversible" equation, the Fokker-Planck equation, yet simple enough to examine both the *forward motion* (i.e., the development in time from uncorrelated initial conditions, a special case of the Green and Piccirelli initial condition) and the corresponding *reverse motion*. The evolution of F_1 for the forward motion is governed by the Fokker-Planck equation while the *corresponding* reverse motion is governed by an anti-Fokker-Planck equation.

II. BASIC EQUATIONS

We start our development with the BBKGY equations, Eq. (1), the first two of which, in the spatially uniform case, are

$$\frac{\partial F_1(\mathbf{p}_1,t)}{\partial t}=\frac{1}{v}\int dx_2\theta_{12}(x_1,x_2)F_2(x_1,x_2,t)\,, \tag{4a}$$

$$\frac{\partial F_2(x_1,x_2,t)}{\partial t}+\mathcal{H}_2(x_1,x_2)F_2(x_1,x_2,t) =\frac{1}{v}\int dx_3(\theta_{13}+\theta_{23})F_3(x_1,x_2,x_3,t)\,. \tag{4b}$$

We introduce the two particle, three particle—correlation functions g and h—by

$$F_2(x_1,x_2,t)=F_1(\mathbf{p}_1,t)F_1(\mathbf{p}_2,t)+g(x_1,x_2,t)\,, \tag{5a}$$

$$F_3(x_1,x_2,x_3,t)=F_1(\mathbf{p}_1,t)F_1(\mathbf{p}_2,t)F_1(\mathbf{p}_3,t)+F_1(\mathbf{p}_1,t)g(x_2,x_3,t)+F_1(\mathbf{p}_2,t)g(x_1,x_3,t)+F_1(\mathbf{p}_3,t)g(x_1,x_2,t)+h(x_1,x_2,x_3,t)\,, \tag{5b}$$

etc. With these definitions, Eqs. (4a) and (4b) are

$$\frac{\partial F_1(\mathbf{p}_1,t)}{\partial t}=\frac{1}{v}\int dx_2\theta_{12}g(x_1,x_2,t)\,, \tag{6a}$$

$$\frac{\partial g(x_1,x_2,t)}{\partial t}+[\mathcal{H}_1(x_1)+\mathcal{H}_1(x_2)]g(x_1,x_2,t) =\theta_{12}[F_1(\mathbf{p}_1,t)F_1(\mathbf{p}_2,t)+g(x_1,x_2,t)]+\frac{1}{v}\int dx_3\theta_{13}F_1(\mathbf{p}_1,t)g(x_2,x_3,t) +\frac{1}{v}\int dx_3\theta_{23}F_1(\mathbf{p}_2,t)g(x_1,x_3,t)+\frac{1}{v}\int dx_3(\theta_{13}+\theta_{23})h(x_1,x_2,x_3,t)\,. \tag{6b}$$

Define the operator $S_{-t}(x_1,x_2)$ in terms of the single-particle Liouville operator $\mathcal{H}_1$ by

$$S_{-t}(x_1,x_2)\equiv\exp\{-t[\mathcal{H}_1(x_1)+\mathcal{H}_1(x_2)]\}\,, \tag{7}$$

i.e., $S_{-t}(x_1,x_2)$ when operating on the point x_i ($i=1, 2$) propagates the point back in time along a *free-particle orbit* over a time interval t. Equation (6b) can be formally integrated along free-particle orbits to give

$$g(x_1,x_2,t)=S_{-(t-t_0)}(x_1,x_2)g(x_1,x_2,t_0)+\int_0^{t-t_0}dt'S_{-t'}(x_1,x_2)\Big\{\theta_{12}[F_1(\mathbf{p}_1,\,t-t')F_1(\mathbf{p}_2,\,t-t')+g(x_1,\,x_2,\,t-t')] +\frac{1}{v}\int dx_3\theta_{13}F_1(\mathbf{p}_1,\,t-t')g(x_2,\,x_3,\,t-t')+\frac{1}{v}\int dx_3\theta_{23}F_1(\mathbf{p}_2,\,t-t')g(x_1,\,x_3,\,t-t') +\frac{1}{v}\int dx_3(\theta_{13}+\theta_{23})h(x_1,\,x_2,\,x_3,\,t-t')\Big\}\,. \tag{8}$$

At this point we introduce the weak short-range potential and assume we can find a solution in the form

$$g(x_1,x_2,t)=g^0(x_1,x_2,t)+\epsilon g^{(1)}(x_1,x_2,t)+\cdots\,, \tag{9}$$
$$h(x_1,x_2,x_3,t)=h^0(x_1,x_2,x_3,t)+\epsilon h^{(1)}(x_1,x_2,x_3,t)+\cdots\,,$$

etc., where ϵ is an expansion parameter that measures the strength of the potential (actually a measure of the ratio of the average potential to the average relative kinetic energy). For our purpose it will be sufficient to consider

probable form for F_s. This has been considered by Grad (Ref. 6) who has shown that most likely the F_s are uncorrelated.

[6] H. Grad, J. Chem. Phys. 33, 1342 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$g^0(x_1,x_2,t_0)=h^0(x_1,x_2,x_3,t_0)=\cdots=0$. The zeroth- and first-order equations are then (dropping the ϵ notation)

$$g^0(x_1,x_2,t)=0\,,$$
$$g^{(1)}(x_1,x_2,t)=S_{-(t-t_0)}(x_1,x_2)g^{(1)}(x_1,x_2,t_0)+\int_0^{t-t_0}dt'S_{-t'}(x_1,x_2)\theta_{12}(x_1x_2)F_1(\mathbf{p}_1,t-t')F_1(\mathbf{p}_2,t-t')\,, \tag{10}$$

etc.

To this order, the equation for F_1 [Eq. (6a)] is

$$\frac{\partial F_1(\mathbf{p}_1,t)}{\partial t}=\frac{1}{v}\int dx_2\theta_{12}g^{(1)}(x_1,x_2,t)\,. \tag{11}$$

In Appendix A we prove that the set of Eqs. (10) and (11) are still reversible; the ϵ expansion has not destroyed the reversibility. This point is taken up again in Sec. IV.

III. RELAXATION OF THE CORRELATION FUNCTION

We now consider a special case, namely (taking $t_0=0$), $g^{(1)}(x_1,x_2,0)=0$.

(a) Consider first that $|\mathbf{q}_1-\mathbf{q}_2|<R$, where R is the range of the force. If we assume there exists a time τ, the "time of a collision," such that for essentially all relevant momenta

$$S_{-\tau}(x_1,x_2)\frac{\partial\varphi_{12}(|\mathbf{q}_1-\mathbf{q}_2|)}{\partial\mathbf{q}_1}$$
$$=\partial\varphi_{12}\left(\left|\mathbf{q}_1-\mathbf{q}_2-\frac{(\mathbf{p}_1-\mathbf{p}_2)}{m}\tau\right|\right)\Big/\partial\mathbf{q}_1=0\,,$$

then Eq. (10) reduces, for $t\gtrsim\tau$ to

$$g^{(1)}(x_1,x_2,t)$$
$$=\int_0^{\tau}dt'S_{-t'}(x_1,x_2)\theta_{12}F_1(\mathbf{p}_1,t-t')F_1(\mathbf{p}_2,t-t')\,,\qquad \text{for}\quad t\gtrsim\tau\,; \tag{12}$$

and, from (11) we have

$$F_1(\mathbf{p}_i,t-t')=F_1(\mathbf{p}_i,t)+O(\tau/\mathcal{T})\ \text{ for }\ 0\leqslant t'\leqslant\tau\,, \tag{13}$$

where $\mathcal{T}$ is the mean time between collisions.

To lowest order,

$$g_{\mathrm{FP}}{}^{(1)}(x_1,x_2,t)$$
$$=\int_0^{\tau}dt'S_{-t'}(x_1,x_2)\theta_{12}F_1(\mathbf{p}_1,t)F_1(\mathbf{p}_2,t)\,,\qquad \text{for}\quad t\gtrsim\tau\,. \tag{14}$$

Therefore, to this order, the functional form proposed by Bogoliubov develops, for $|\mathbf{q}_1-\mathbf{q}_2|<R$, in a time τ from the initial time. The subscript F-P on $g^{(1)}$ indicates that this is the form of $g^{(1)}$ that together with Eq. (11) is the Fokker-Planck equation (or, as Bogoliubov calls it, the Landau equation). The Fokker-Planck equation is irreversible in the sense that it obeys an H theorem.[7]

[7] A. Lenard, Ann Phys. (N. Y.) 10, 390 (1960).

(b) Consider $|\mathbf{q}_1-\mathbf{q}_2|>R$.

Suppose $\mathbf{p}_1$ and $\mathbf{p}_2$ are such that at no earlier time have the two points interacted, i.e., $|S_{-t}(x_1,x_2)\{\mathbf{q}_1-\mathbf{q}_2\}|>R$ for $0\leqslant t\leqslant\infty$; then $g^{(1)}(x_1,x_2,t)=0$ for all time.

If $\mathbf{p}_1$ and $\mathbf{p}_2$ are such that at an earlier time $\tau_0(x_1,x_2)$ the points are just starting to interact, i.e.,

$$|S_{-\tau_0}\{\mathbf{q}_1-\mathbf{q}_2\}|=R\,,$$

then Eq. (10) is

$$g^{(1)}(x_1,x_2,t)$$
$$=0\,,\quad t\leqslant\tau_0\,;$$
$$=\int_{\tau_0}^{t}dt'S_{-t'}(x_1x_2)\theta_{12}F_1(\mathbf{p}_1,t-t')F_1(\mathbf{p}_2,t-t')\,,\qquad \tau_0\leqslant t\leqslant\tau_0+\tau\,; \tag{15}$$
$$=\int_{\tau_0}^{\tau_0+\tau}dt'S_{-t'}\theta_{12}F_1(\mathbf{p}_1,t-t')F_1(\mathbf{p}_2,t-t')\,,\qquad t\gtrsim\tau_0+\tau\,.$$

The asymptotic form for $g^{(1)}$ is reached only after $t\gtrsim\tau_0+\tau$ which can be (for given $\mathbf{p}_1$ and $\mathbf{p}_2$) arbitrarily large for points $\mathbf{q}_1$ and $\mathbf{q}_2$ sufficiently far apart. It is interesting to notice that the previous history of the system, in terms of F_1, is stored in the correlation function of distant points; for, at any time t, there are pairs of points $(\mathbf{q}_1,\mathbf{q}_2)$ such that the corresponding $g^{(1)}$ depends on the value of F_1 at any given time between zero and t.

The asymptotic form of (15) can be written as

$$g^{(1)}(x_1,x_2,t)=\mathbf{J}(t-\tau_0)\cdot\mathbf{I}_{r_0}\,,\quad t\gtrsim\tau_0+\tau\,,$$
$$\mathbf{J}(t)=\left(\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right)F_1(\mathbf{p}_1,t)F_1(\mathbf{p}_2,t)\,, \tag{16}$$
$$\mathbf{I}_{r_0}=\int_{\tau_0}^{\tau_0+\tau}dt'S_{-t'}(x_1,x_2)\frac{\partial\varphi_{12}}{\partial\mathbf{q}_1}\,,$$

where we have again used Eq. (13) and, from the spherical symmetry of the force law and the fact that the collision is completed, $\mathbf{I}_{r_0}$ is in the direction of the distance of closest approach $\mathbf{r}_0$. Therefore, $g^{(1)}$ is nonzero if $\mathbf{J}$ has a component in the $\mathbf{r}_0$ direction. In equilibrium, $\mathbf{J}$ is in the direction of the relative momentum $\mathbf{P}=\mathbf{p}_2-\mathbf{p}_1$ which is normal to $\mathbf{r}_0$ so that $g^{(1)}=0$. We see then, in the nonequilibrium state, the range of $g^{(1)}$ can be arbitrarily large, but the range of $(\mathbf{p}_1,\mathbf{p}_2)$ which gives a nonzero value to $g^{(1)}$ goes down with increasing $|\mathbf{q}_1-\mathbf{q}_2|$. In Appendix B we show that the long range

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of $g^{(1)}$ does not cause the next approximation $g^{(2)}$ to become unbounded with time.

IV. THE REVERSE MOTION

In Sec. IIIa we considered the evolution of a system that at zero time was uncorrelated. In this section we shall describe the corresponding reverse motion. For this purpose we consider the evolution of a second system constructed at an arbitrary time $t_0>0$ that is identical to the first system but with all momenta reversed. This second system is initially correlated in a special way and it is this initial correlation which causes the reverse motion to be accomplished.

Consider again the initial condition $g^{(1)}(x_1,x_2,0)=0$. We have, from Eq. (10) with $t_0=0$

$$g^{(1)}(0|x_1,x_2,t) = \int_0^t dt' S_{-t'}(x_1,x_2)\theta_{12}F_1(0|\mathbf{p}_1,t-t')\times F_1(0|\mathbf{p}_2,t-t'), \quad (17)$$

where the zero preceding the vertical bar in $g^{(1)}$ and F_1 indicate the initial time was zero. In particular (suppressing $\mathbf{q}_1$, $\mathbf{q}_2$) we have

$$g^{(1)}(0|-\mathbf{p}_1,-\mathbf{p}_2,t) = \int_0^t dt' S_{-t'}(-\mathbf{p}_1,-\mathbf{p}_2)\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)\times F_1(0|-\mathbf{p}_1,t-t')F_1(0|-\mathbf{p}_2,t-t'). \quad (18)$$

At time t_0 start another system off with initial conditions

$$F_1(t_0|\mathbf{p}_1,t)|_{t=t_0}=F_1(0|-\mathbf{p}_1,t_0), \quad (19a)$$

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)|_{t=t_0}=g^{(1)}(0|-\mathbf{p}_1,-\mathbf{p}_2,t_0). \quad (19b)$$

For $t\geq t_0$ we have from Eq. (10)

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t) = S_{-(t-t_0)}(\mathbf{p}_1,\mathbf{p}_2)g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t_0) + \int_0^{t-t_0} dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)\times F_1(t_0|\mathbf{p}_1,t-t')F_1(t_0|\mathbf{p}_2,t-t'). \quad (20)$$

Using (19b) and (18) in (20) we have

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)=S_{-(t-t_0)}(\mathbf{p}_1,\mathbf{p}_2)\left\{\int_0^{t_0} dt' S_{-t'}(-\mathbf{p}_1,-\mathbf{p}_2)\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)F_1(0|-\mathbf{p}_1,t_0-t')F_1(0|-\mathbf{p}_2,t_0-t')\right\} + \int_0^{t-t_0} dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)F_1(t_0|\mathbf{p}_1,t-t')F_1(t_0|\mathbf{p}_2,t-t'). \quad (21)$$

In the first term in Eq. (21) we can combine the S operators

$$S_{-(t-t_0)}(\mathbf{p}_1,\mathbf{p}_2)S_{-t'}(-\mathbf{p}_1,-\mathbf{p}_2)=S_{-(t-t_0-t')}(\mathbf{p}_1,\mathbf{p}_2)=S_{-(t'+t_0-t)}(-\mathbf{p}_1,-\mathbf{p}_2). \quad (22)$$

Splitting the range of integration in the first integral in (21) into

$$(0,t-t_0) \quad \text{and} \quad (t-t_0,t_0) \quad \text{for} \quad t_0\leq t\leq 2t_0,$$

the first integral is

$$\left\{\int_0^{t-t_0} dt' S_{-(t-t_0-t')}(\mathbf{p}_1,\mathbf{p}_2)+\int_{t-t_0}^{t_0} dt' S_{-(t'+t_0-t)}(-\mathbf{p}_1,-\mathbf{p}_2)\right\}\theta_{12}F_1F_1.$$

Changing time variables, Eq. (21) becomes

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)=\int_0^{t-t_0} dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)F_1(0|-\mathbf{p}_1,2t_0+t'-t)F_1(0|-\mathbf{p}_2,2t_0+t'-t) + \int_0^{2t_0-t} dt' S_{-t'}(-\mathbf{p}_1,-\mathbf{p}_2)\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)F_1(0|-\mathbf{p}_1,2t_0-t'-t)F_1(0|-\mathbf{p}_2,2t_0-t'-t) + \int_0^{t-t_0} dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)F_1(t_0|\mathbf{p}_1,t-t')F_1(t_0|\mathbf{p}_2,t-t'), \quad \text{for} \quad t_0\leq t\leq 2t_0. \quad (23)$$

From the general reversibility argument (Appendix A) we would expect

$$F_1(t_0|\mathbf{p}_1,t)=F_1(0|-\mathbf{p}_1,2t_0-t), \quad (24a)$$

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)=g^{(1)}(0|-\mathbf{p}_1,-\mathbf{p}_2,2t_0-t), \quad \text{for} \quad t_0\leq t\leq 2t_0, \quad (24b)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which satisfy the initial conditions (19a) and (19b). To show that this is the case, suppose (24a) is the solution of (23) and (11), then the first term in (23) cancels the last term and the remaining term is, by (18), $g^{(1)}(0|-\mathbf{p}_1, -\mathbf{p}_2, 2t_0-t)$ thus establishing (24b); and (24b) in (11) leads to (24a), i.e., Eq. (11) is

$$\frac{\partial F_1(t_0|\mathbf{p}_1,t)}{\partial t}=-\frac{1}{v}\int dx_2\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)\,. \tag{25}$$

The right-hand side of (25) is

$$-\frac{1}{v}\int dx_2\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)g^{(1)}(0|-\mathbf{p}_1,-\mathbf{p}_2,2t_0-t)=-\frac{1}{v}\int dx_2\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)g^{(1)}(0|-\mathbf{p}_1,-\mathbf{p}_2,2t_0-t) \tag{25a}$$

and the right side of (25a) is, by Eq. (11), $\partial F_1(0|-\mathbf{p}_1, 2t_0-t)/\partial t$; we have for $t_0\leq t\leq 2t_0$, $\partial F_1(t_0|\mathbf{p}_1,t)/\partial t = \partial F_1(0|-\mathbf{p}_1, 2t_0-t)/\partial t$, or $F_1(t_0|\mathbf{p}_1,t)-F_1(t_0|\mathbf{p}_1,t_0)=F_1(0|-\mathbf{p}_1, 2t_0-t)-F_1(0|-\mathbf{p}_1, t_0)$ and, with initial condition (19a), this is (24a).

Over the time interval $t_0\leq t\leq 2t_0$, the equation for $F_1(t_0|\mathbf{p}_1,t)$ can be obtained from (25) and (23). Equation (23) can be written, using (24a) and (13), as

$$\begin{aligned} g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,t)&=\int_0^{2t_0-t}dt'S_{-t'}(-\mathbf{p}_1,-\mathbf{p}_2)\theta_{12}(-\mathbf{p}_1,-\mathbf{p}_2)F_1(0|-\mathbf{p}_1,2t_0-t'-t)F_1(0|-\mathbf{p}_2,2t_0-t'-t)\\ &=-\int_0^{2t_0-t}dt'S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)F_1(t_0|\mathbf{p}_1t)F_1(t_0|\mathbf{p}_2,t)\,. \end{aligned} \tag{26}$$

Equation (26), together with (25), is the anti-Fokker-Planck equation, i.e., the collision term is the negative of the Fokker-Planck collision term. If we call $\mathbf{q}_1-\mathbf{q}_2=\mathbf{r}$ we have

$$\partial F_1(t_0|\mathbf{p}_1,t)/\partial t=-\frac{1}{v}\int d\mathbf{p}_2d\mathbf{r}\left[\frac{\partial\varphi(|\mathbf{r}|)}{\partial\mathbf{r}}\cdot\left\{\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right\}\right]\int_0^{2t_0-t}dt'\partial\varphi\left(\left|\mathbf{r}+\frac{(\mathbf{p}_1-\mathbf{p}_2)}{m}t'\right|\right)\Big/\partial\mathbf{r}$$
$$\cdot\left\{\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right\}F_1(t_0|\mathbf{p}_1,t)F_1(t_0|\mathbf{p}_2,t)\,.$$

Changing $\mathbf{r}\rightarrow-\mathbf{r}$

$$\frac{\partial F_1(t_0|\mathbf{p}_1,t)}{\partial t}=-\frac{1}{v}\int d\mathbf{p}_2d\mathbf{r}\left[\frac{\partial\varphi(|\mathbf{r}|)}{\partial\mathbf{r}}\cdot\left\{\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right\}\right]\int_0^{2t_0-t}dt'\partial\varphi\left(\left|\mathbf{r}-\frac{(\mathbf{p}_1-\mathbf{p}_2)}{m}t'\right|\right)\Big/\partial\mathbf{r}$$
$$\cdot\left\{\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right\}F_1(t_0|\mathbf{p}_1,t)F_1(t_0|\mathbf{p}_2,t)\,, \tag{27}$$

which is the anti-Fokker-Planck equation for $t_0\leqslant t\leqslant 2t_0-\tau$.

At time $t=2t_0$, from (24a) and (24b) we obtain

$$g^{(1)}(t_0|\mathbf{p}_1,\mathbf{p}_2,2t_0)=0\,, \tag{28a}$$

$$F_1(t_0|\mathbf{p}_1,2t_0)=F_1(0|-\mathbf{p}_1,0)\,. \tag{28b}$$

The equation describing the motion for $t\geqslant 2t_0$ is then, from (11) and (10)

$$\frac{\partial F_1(2t_0|\mathbf{p}_1,t)}{\partial t}=-\frac{1}{v}\int dx_2\theta_{12}g^{(1)}(2t_0|\mathbf{p}_1,\mathbf{p}_2,t)\,, \tag{29a}$$

$$g^{(1)}(2t_0|\mathbf{p}_1,\mathbf{p}_2,t)$$
$$=\int_0^{t-2t_0}dt'S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\theta_{12}(\mathbf{p}_1,\mathbf{p}_2)$$
$$\times F_1(2t_0|\mathbf{p}_1,t)F_1(2t_0|\mathbf{p}_2,t)\,, \tag{29b}$$

where

$$g^{(1)}(2t_0|\mathbf{p}_1,\mathbf{p}_2,2t_0)=0\,,$$
$$F_1(2t_0|\mathbf{p}_1,t)|_{t=2t_0}=F_1(0|-\mathbf{p}_1,0)\,, \tag{29c}$$

and, therefore, after $t\geqslant 2t_0+\tau$ the equation for F_1 is again the Fokker-Planck equation and the system is again approaching equilibrium.

V. SUMMARY

From the BBKGY set of equations, we have derived an approximate set of equations for the weak force case, Eqs. (10) and (11); these have been shown to be reversible in Appendix A. In the special case that the initial $g^{(1)}=0$, the equation for F_1 evolves into the "irreversible" Fokker-Planck equation after a time τ (on the order of a time of a collision) and the system approaches equilibrium.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Suppose, on the other hand, the system is allowed to run for an arbitrary length of time t_0; then in order to describe the reverse motion, we consider a second system identical to the first at time t_0, but with all momenta reversed. The initial conditions on F_1 and $g^{(1)}$ for this second system are constructed from the F_1 and $g^{(1)}$ of the first system at t_0. The second system during the time interval $(t_0,2t_0)$ now performs exactly the reverse motion that the original system performed during the time $(0,t_0)$. The equation for F_1 for the second system during the time interval $(t_0, 2t_0-\tau)$ is the anti-Fokker-Planck equation and after the time $2t_0+\tau$ is the Fokker-Planck equation.

For the second system the dependence of $g^{(1)}$ on its initial value causes the reverse motion to be accomplished: "built into" the initial value of $g^{(1)}$ at pairs of points outside the range of the force is the earlier behavior of F_1 for the first system. The dependence of $g^{(1)}$ on its initial value would be missing if the momentum reversal were applied to the Fokker-Planck equation.

The final question of what to use in describing a given system is answered in the spirit of statistical mechanics by the argument that if all that is known initially about a system is F_1, then the most likely $g^{(1)}$ is zero and F_1 will most likely evolve according to the Fokker-Planck equation.[5]

APPENDIX A

Equations (10) and (11), with $t_0=0$, suppressing spatial coordinates in $g^{(1)}$ and S, are

$$\frac{\partial F_1(\mathbf{p}_1,t)}{\partial t}=-\frac{1}{v}\int d\mathbf{q}_2 d\mathbf{p}_2 \frac{\partial\varphi_{12}}{\partial\mathbf{q}_1}\cdot\left(\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right)g^{(1)}(\mathbf{p}_1,\mathbf{p}_2,t), \quad \text{(A1)}$$

$$g^{(1)}(\mathbf{p}_1,\mathbf{p}_2,t) = S_{-t}(\mathbf{p}_1,\mathbf{p}_2)g^{(1)}(\mathbf{p}_1,\mathbf{p}_2,0) + \int_0^t dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\frac{\partial\varphi_{12}}{\partial\mathbf{q}_1}\cdot\left(\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right) \times F_1(\mathbf{p}_1,t-t')F_1(\mathbf{p}_2,t-t'). \quad \text{(A2)}$$

Consider $\mathfrak{F}(\mathbf{p}_1,t)$ and $\mathfrak{G}(\mathbf{p}_1,\mathbf{p}_2,t)$ to be the solutions of (A1) and (A2) with initial conditions $\mathfrak{F}(\mathbf{p}_1,0)$ and $\mathfrak{G}(\mathbf{p}_1,\mathbf{p}_2,0)$. In Eqs. (A1) and (A2) change $\mathbf{p}_i\rightarrow-\mathbf{p}_i$ and $t\rightarrow-t$; then using the fact that $S_{-t}(\mathbf{p}_1,\mathbf{p}_2)=S_t(-\mathbf{p}_1,-\mathbf{p}_2)$ we have

$$\frac{\partial F_1(-\mathbf{p}_1,-t)}{\partial t}=-\frac{1}{v}\int d\mathbf{q}_2 d\mathbf{p}_2 \frac{\partial\varphi_{12}}{\partial\mathbf{q}_1}\cdot\left(\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right) \times g^{(1)}(-\mathbf{p}_1,-\mathbf{p}_2,-t), \quad \text{(A3)}$$

$$g^{(1)}(-\mathbf{p}_1,-\mathbf{p}_2,-t) = S_{-t}(\mathbf{p}_1,\mathbf{p}_2)g^{(1)}(-\mathbf{p}_1,-\mathbf{p}_2,0) + \int_0^t dt' S_{-t'}(\mathbf{p}_1,\mathbf{p}_2)\frac{\partial\varphi_{12}}{\partial\mathbf{q}_1}\cdot\left(\frac{\partial}{\partial\mathbf{p}_1}-\frac{\partial}{\partial\mathbf{p}_2}\right) \times F_1(-\mathbf{p}_1,-t+t')F_1(-\mathbf{p}_2,-t+t'). \quad \text{(A4)}$$

Comparing Eqs. (A3) and (A4) with (A1) and (A2) we see that if we take as initial conditions in (A3) and (A4)

$$F_1(-\mathbf{p}_1,-t)\big|_{t=0}=\mathfrak{F}(\mathbf{p}_1,0),$$
$$g^{(1)}(-\mathbf{p}_1,-\mathbf{p}_2,-t)\big|_{t=0}=\mathfrak{G}(\mathbf{p}_1,\mathbf{p}_2,0),$$

then the solutions of (A3) and (A4) are

$$F_1(-\mathbf{p}_1,-t)=\mathfrak{F}(\mathbf{p}_1,t),$$
$$g^{(1)}(-\mathbf{p}_1,-\mathbf{p}_2,-t)=\mathfrak{G}(\mathbf{p}_1,\mathbf{p}_2,t),$$

or

$$F_1(\mathbf{p}_1,t)=\mathfrak{F}(-\mathbf{p}_1,-t),$$
$$g^{(1)}(\mathbf{p}_1,\mathbf{p}_2,t)=\mathfrak{G}(-\mathbf{p}_1,-\mathbf{p}_2,-t). \quad \text{(A5)}$$

Therefore, the approximation used to obtain Eqs. (A1) and (A2) have not destroyed the reversibility.

APPENDIX B

We show here that the long range of $g^{(1)}$ (see Sec. IIIb) does not cause the next approximation $g^{(2)}$ to become unbounded with increasing time. $g^{(2)}$ will contain a term

$$\frac{1}{v}\int_0^t dt' S_{-t'}(x_1,x_2)\int dx_3\theta_{13}F_1(\mathbf{p}_1,t-t')g^{(1)}(x_2,x_3,t-t') = -\frac{1}{v}\int_0^t dt'\int d\mathbf{q}_3 d\mathbf{p}_3\frac{\partial\varphi_{13}}{\partial\mathbf{q}_1}(|\mathbf{q}_1'-\mathbf{q}_3|)\cdot\frac{\partial F_1(\mathbf{p}_1,t-t')}{\partial\mathbf{p}_1} \times g^{(1)}(x_2',x_3,t-t'), \quad \text{(B1)}$$

where

$$x_i'=S_{-t'}(x_1,x_2)x_i; \quad i=1,2.$$

Figure 1 shows a configuration at time t' where the integrand in (B1) is nonzero.

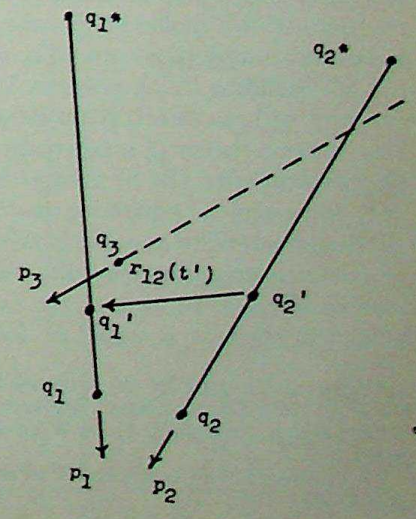


FIG. 1. An interacting configuration. (a) $\mathbf{q}_i^*=\mathbf{q}_i-\mathbf{p}_i t/m$; $i=1,2$. (b) $\mathbf{q}_3$ is within range of force of $\mathbf{q}_1'$ and (c) $\mathbf{p}_3$ is such that at an earlier time (but before particle 2 is at $\mathbf{q}_2^*$) particles 2 and 3 interact since $g^{(1)}\neq0$ at that time.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In terms of the relative momentum $\mathbf{P}_{32}$; $\mathbf{P}_{32}$ must be within a solid angle $\pi R^2/|\mathbf{r}_{23}(t')|^2$ in the direction $\mathbf{r}_{23}(t')$, which at large distances can be replaced by $\pi R^2/|\mathbf{r}_{12}(t')|^2$, where $\mathbf{r}_{12}(t')=\mathbf{r}_{12}-\mathbf{P}_{12}t'/m$ and $\mathbf{r}_{12}=\mathbf{q}_1-\mathbf{q}_2$, $\mathbf{P}_{12}=\mathbf{p}_1-\mathbf{p}_2$.

In (B1) changing the momentum integration to $\mathbf{P}_{32}$ we have

$$\int_0^t dt' \int d\Omega_{P_{32}}\alpha\,, \tag{B2}$$

$$\alpha=-\frac{1}{v}\int d\mathbf{q}_3 dP_{32}P_{32}^2\frac{\partial\varphi_{13}}{\partial\mathbf{q}_1}\cdot\frac{\partial F_1}{\partial\mathbf{p}_1}g^{(1)}.$$

To obtain the magnitude of (B2) we neglect the variation of α with time and treat it as constant with respect to the angles when it is nonzero. We then have

$$\alpha\int_0^t dt'\frac{\pi R^2}{|\mathbf{r}_{12}(t')|^2}$$
$$=\frac{\alpha\pi R^2}{r_{12}(P_{12}/m)(1-\cos^2\theta)^{1/2}}\left\{\tan^{-1}\left[\frac{P_{12}t/mr_{12}-\cos\theta}{(1-\cos^2\theta)^{1/2}}\right]-\tan^{-1}\left[\frac{-\cos\theta}{(1-\cos^2\theta)^{1/2}}\right]\right\}, \tag{B3}$$

where θ is the angle between $\mathbf{r}_{12}$ and $\mathbf{P}_{12}$. Equation (B3) approaches a limit as $t\to\infty$, according to

$$\frac{\alpha\pi R^2}{r_{12}(P_{12}/m)(1-\cos^2\theta)^{1/2}}\left\{\frac{\pi}{2}-\tan^{-1}\left[\frac{-\cos\theta}{(1-\cos^2\theta)^{1/2}}\right]-\frac{(1-\cos^2\theta)^{1/2}}{P_{12}t/mr_{12}}\right\}. \tag{B4}$$

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Possibility of Synthesizing an Organic Superconductor*

W. A. Little
Department of Physics, Stanford University, Stanford, California

(Received 13 November 1963; revised manuscript received 27 January 1964)

London's idea that superconductivity might occur in organic macromolecules is examined in the light of the BCS theory of superconductivity. It is shown that the criterion for the occurrance of such a state can be met in certain organic polymers. A particular example is considered in detail. From a realistic estimation of the matrix elements and density of states in this polymer it is concluded that superconductivity should occur even at temperatures well above room temperature. The physical reason for this remarkable high transition temperature is discussed. It is shown further that the superconducting state of these polymers should be distinguished by certain unique chemical properties which could have considerable biological significance.

I. INTRODUCTION

IN the forward to Vol. 1 of his monographs on superfluids, F. London[1] questions whether a superfluid-like state might occur in certain macromolecules which play an important role in biochemical reactions. If this should be the case, an entirely new and important consideration would be added to the problem of understanding living systems. In view of the significance of such an effect, it appears appropriate at this time, when a theory of superconductivity, the Bardeen-Cooper-Schrieffer (BCS) theory[2] has been so remarkably successful in explaining much of the behavior of superconductors, to examine in the light of this whether or not a superconducting state might occur in certain macromolecules. In view of the extreme complexity of biological systems, it would be folly for a physicist to attempt to experiment in such an environment. Instead of attempting this, we shall tackle the problem on our own grounds. The BCS theory, while by no means complete and exact, has succeeded in providing a model with most of the essential features of a superconductor. In particular, it prescribes certain criteria for a system which, if satisfied, should lead to the superconducting state. Our approach is to consider how these criteria might be applied to the design of a particular organic molecule which, if its synthesis is possible, should show some of the essential features of a superconductor and, as we shall show, some remarkable chemical properties as well. One of the interesting features about the particular class of molecules we investigate in detail is that the molecules should be superconducting at room temperature and, indeed, to temperatures well above room temperatures. We can show on simple physical grounds why this is so and perhaps, with hindsight, why this was to be expected.

The idea of superconductivity in organic systems is not a new idea, however, there is a considerable amount

* Supported in part by the National Science Foundation and the U. S. Navy Office of Naval Research.

[1] F. London, *Superfluids* (John Wiley & Sons, Inc., New York, 1950), Vol. 1.

[2] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of confusion as to the exact meaning of this. The diamagnetic ring currents of aromatic molecules such as benzene, naphthalene, etc., are nondissipative currents similar in many respects to the persistent currents of superconducting rings and, have often been referred to as a form of superconductivity. However, the "superconductivity" of these molecules is not the same as the superconductivity of bulk materials. The reason, I believe, is the following. In macroscopically large superconductors, if superconductivity exists, then a finite fraction of the charge carriers, in general, the BCS pairs are in identically the same center-of-mass momentum state. This state then has a macroscopic occupation. In a magnetic field the canonical momentum of this state remains unchanged, but due to the vector potential term contained in it a current is induced and the energy of the state changes. For a macroscopically large superconductor the kinetic energy of the different center-of-mass momentum states of the pairs lie extremely close to one another, however, because the coherence energy of each state depends upon the square of the number of pairs in that state, the state which is macroscopically occupied is appreciably lower in energy than any of the neighboring states even in a moderate magnetic field. It is only by transitions in which practically all the pairs in the macroscopically occupied state simultaneously move to another state that a lower energy final state can be reached. This is obviously highly forbidden and, consequently, the system of pairs remains in the momentum state into which condensation originally occurred. Thus, it is the coherence energy which prevents the system from freely adjusting itself to take the lowest possible energy. In the aromatic ring compounds practically all the molecules are in their ground states. In a magnetic field the canonical momentum of the electrons in this state remain unchanged and diamagnetic currents flow in the molecule similar to those of a bulk superconductor. The energy of the different momentum states of the electrons in each molecule in this case are well separated though, because the molecules are of microscopic size. Thus, the momenta of the electrons do not change because for fields as large as those available in the laboratory, the state which evolves out of the original ground state still is lower in energy than any other in the presence of the field. If, however, the aromatic system is made arbitrarily large such as in graphite, bulk superconductivity does not result because as the system gets bigger, the different momentum states of the electrons approach each other in energy. Transitions can then occur between states and the induced currents are dissipated. So that in order to get superconductivity in a macromolecule or in a bulk material, something of the nature of a coherence energy is required. In conventional superconductors this is provided by the phonon-induced, electron-electron interaction; in attempting to devise a macromolecule which is to be superconducting one must provide, therefore, some mechanism similar to this. In our model we do this in the following manner.

II. MODEL SYSTEM

We shall consider a molecule consisting of two parts, a long chain called the "spine" in which electrons fill the various states and may or may not form a conducting system; and secondly, a series of arms or side chains attached to the spine as indicated in Fig. 1. We will show that by appropriate choice of the molecules which constitute the side chains, the virtual oscillation of charge in these side chains can provide an interaction between the electrons moving in the spine. This can be made a sufficiently attractive interaction so that the superconducting state results. We can show further that even if the spine by itself is initially an insulator due to the valence band being full and the conduction band empty, the addition of side chains can increase the electron-electron attraction to the point where it becomes energetically favorable to enter the superconducting state by mixing in states of the conduction band. The spine thus transforms from the insulating or semiconducting state directly to the superconducting metallic state upon the addition of the side chains.

Consider a long chain molecule as shown in the left half of Fig. 1. We will assume this is a conjugated chain of double and single bonds resonating between the two at each link. This corresponds in the band theory of metals to a band which is half filled and ideally is a metallic conductor. (See, however, Sec. III.) At the points P, P', $\cdots$, a regular array of side chain molecules B are attached. The individual side-chain molecules are chosen to have a low-lying excited state such that transitions from the ground state to the excited state correspond classically to an oscillation of charge from end to end of the molecule.

The electrons moving in the spine may be described

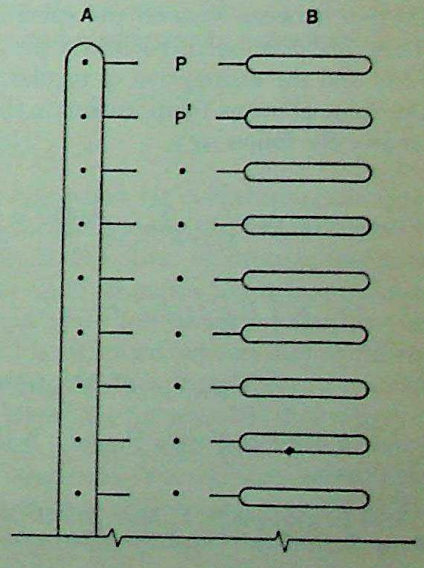


FIG. 1. Proposed model of a superconducting organic molecule. The molecule A is a long unsaturated polyene chain called the "spine." The molecules B are side chains attached to the spine at points P, P', ···.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in the tight-binding approximation[3] by eigenfunctions of the form

$$\phi_{k,m}(r)=\frac{1}{G^{1/2}}\sum_i e^{ikR_i}U_m(r-R_i)\,, \qquad (2.1)$$

where $U_m(r-R_i)$ is the wave function of an electron in a single atom located at R_i. G is the number of links in the chain and ϵ_k is the energy of this electron and we will assume cyclic boundary conditions. To avoid unnecessary complication we shall ignore the electron spin throughout the discussion except where necessary.

The wave function of the electrons in a single side-chain molecule in an eigenstate n we designate by $\psi_n(r_1,r_2,\cdots)$. Due to the interaction of one side chain upon its neighbors, the degeneracy of the levels of the group of side chains will be removed when the side chains are brought together as in Fig. 1. The band of levels of the system of side chains as a whole can be described then by a new wave function similar to (2.1) above.

$$X_{q,n}=\frac{1}{G^{1/2}}\sum_j e^{iqR_j}\psi_n(r_1-R_j;r_2-R_j\cdots)\,. \qquad (2.2)$$

The Coulomb interaction between the electrons in the spine and those on the side chain will provide an interaction between the electrons in the spine and the side-chain modes. Let this Coulomb interaction be $V_1=\sum_{i,j}V(r-R_i;r_1-R_j,r_2-R_j,\cdots)$ which will give rise to a typical interaction matrix element

$$\int\int\cdots\int\phi_{k'm'}{}^*(r)X_{q'm'}{}^*(r_1,r_2\cdots)$$
$$\times V_1X_{qn}(r_1,r_2\cdots)\phi_{km}(r)d^3rd^3r_1\cdots\,. \qquad (2.3)$$

In the model we shall consider, screening reduces the range of the Coulomb interaction so that one can safely disregard the contribution to the interaction from any but the neighboring side chains. Secondly, in our model there is negligible overlap between side chains, and also overlap between sites on the spine is relatively small. These features together with the properties of (2.1) and (2.2) and the assumption of regular substitution along the spine allow us to approximate the interaction (2.3) to give the following

$$V_{k'm',q'n',qn,km}=\frac{1}{G}\sum_{n''=-1,0,+1}\cos Qn''a\int\cdots$$
$$\times\int U_{m'}{}^*(r)\psi_{n'}{}^*(r_1-n''a,\cdots)$$
$$\times V_1\psi_n(r_1-n''a)U_m(r)\times d^3rd^3r_1\cdots\,, \qquad (2.4)$$

where $k'=k+Q$, $q'=q-Q$, and a is the spacing between side chains.

[3] See for example N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Clarendon Press, Oxford, 1936).

It is convenient to write this as $(1/G)V(Q)$ for the moment. Then the complete spine side-chain interaction can be described in the representation of second quantization as

$$V_{\text{side chain}}=\frac{1}{G}\sum_{\substack{Q,n,n'\\k,m,m'}}V(Q)\Big(\sum_q c_{q-Q,n'}{}^\dagger c_{q,n}\Big)a_{k+Q,m'}{}^\dagger a_{km}\,, \qquad (2.5)$$

where the $a^\dagger$ and a are the creation and destruction operators for the electron in the spine, and $c^\dagger$ and c the corresponding operators for the side-chain modes. The particle-hole operator $\sum_q c_{q-Q,n'}{}^\dagger c_{qn}$ which occurs in (2.5) appears in a similar manner to the phonon creation operator $b_Q{}^\dagger$ in the phonon-electron interaction in a metal. The terms linear in this particle-hole operator which appear through the interaction (2.5) in the total Hamiltonian can be eliminated by the same type of transformation[4] which eliminates the terms linear in the phonon creation operator $b_Q{}^\dagger$ in the conventional theory. This leads to a side-chain induced electron-electron interaction V_2 *between* electrons in the spine of the form

$$V_2=\tfrac{1}{2}\sum_{k,k',Q}\left\{\frac{2V^2(Q)}{G}\,\frac{\hbar\omega_Q}{(\epsilon_{k+Q}-\epsilon_k)^2-(\hbar\omega_Q)^2}\right\}$$
$$\times a_{k+Q}{}^\dagger a_{k'-Q}{}^\dagger a_{k'}a_k\,. \qquad (2.6)$$

We have written $\hbar\omega$ for the difference $(E_{n'}-E_n)$ between the energy of the states of the side-chain modes and have considered for simplicity only one excited state n'. This is identical in form to the usual phonon-induced electron-electron coupling. We note that if $|\hbar\omega|\gg|(\epsilon_{k+Q}-\epsilon_k)|$ then the term in brackets reduces to an attractive interaction, $V\approx-2\langle|V(Q)|^2\rangle_{\text{av}}/G\hbar\omega$, where $\langle|V(Q)|^2\rangle_{\text{av}}$ is the average of the square of the interaction $V(Q)$.

In addition to this attractive term, the screened Coulomb interaction gives a repulsive term. The sum of these two is the net electron-electron interaction. Let this sum be $\langle V(Q)\rangle_{\text{av}}$; then the total Hamiltonian for the electrons in the spine is

$$\mathcal{H}=\sum_k(\epsilon_k-\mu)a_k{}^\dagger a_k$$
$$+\tfrac{1}{2}\sum_{k,k',Q}\langle V(Q)\rangle_{\text{av}}a_{k+Q}{}^\dagger a_{k'-Q}{}^\dagger a_{k'}a_k\,, \qquad (2.7)$$

where $\langle V(Q)\rangle_{\text{av}}$ is given above and μ is the Fermi energy.

The BCS theory then shows that a superconducting state is possible if the following equation can be satisfied for a nonzero gap, Δ:

$$-1=\sum_k\frac{V\tanh\frac{1}{2}\beta E_k}{2E_k}\,, \qquad (2.8)$$

[4] J. Bardeen, *Encyclopedia of Physics*, edited by S. Flügge (Springer-Verlag, Berlin, 1956), Vol. 15, p. 352.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where $E_k=[(\epsilon_k-\mu)^2+\Delta^2]^{1/2}$ and the sum is limited by $|\epsilon_k-\mu|<\hbar\omega$. V is the average over the region of Q where $\langle V(Q)\rangle_{\rm av}$ is attractive. The critical temperature is given by

$$kT_c=1.14\hbar\omega\exp[-1/N(0)V], \qquad (2.9)$$

where $N(0)$ is the density of states of one spin at the Fermi surface.

It is appropriate to dispose of a difficulty here which will arise shortly in considering a particular example of a molecule. In long-chain conjugated molecules the double and single bonds do not resonate freely between each position in the chain, but tend to localize so that a stationary periodic charge distribution and periodic bond length is established. This introduces a periodic potential of twice the atomic spacing in the chain which in turn produces a gap in the density of states halfway up the band. As there is one electron per atom in the band, this band is filled up to the new gap. The conjugated system is then no longer a metallic conductor but a semiconductor or insulator, i.e., the density of states at the Fermi surface $N(0)$ would be zero.[5] However, it is incorrect to interpret Eq. (2.9) as indicating that $T_c=0$ in this case, for Eq. (2.9) is only an approximation. Instead, one must go back to (2.8) and examine this to see whether in this case a solution is possible. In the next section we consider this in detail.

III. SUPERCONDUCTIVITY IN A SEMICONDUCTOR

Let us consider a somewhat idealized case of a semiconductor with a band structure as shown in Fig. 2. Before the introduction of the periodic potential which generates the gap, the energy of the states would have been given by that shown by the dashed line in the figure. Let us take the band gap as δ and for convenience we will assume the density of states near the band edge is the same in the two bands. If the lower band is completely filled and the upper empty, then the Fermi level will be halfway between the two bands. We will assume that $\hbar\omega$ is greater than δ.

Let $N(0)$ be the density of states at the Fermi surface for the system prior to the introduction of the periodic potential and let ϵ be the energy of these states measured with respect to the Fermi surface. The semiconductor band gap can then be conveniently introduced by changing the energy of each state ϵ to $\epsilon'=[\epsilon^2+(\delta/2)^2]^{1/2}$. We can now use Eq. (2.8) to see if the superconducting state can occur at any temperature. At $T=0°$K a

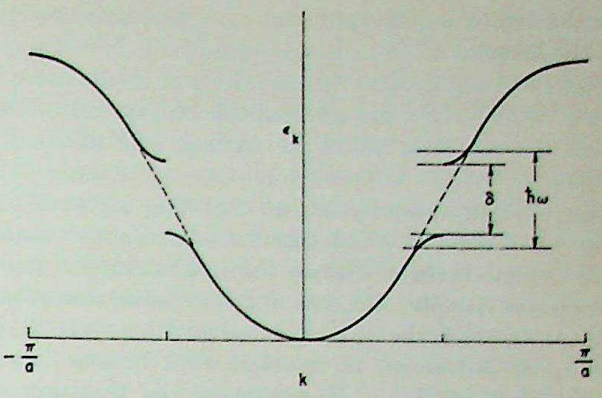


FIG. 2. Band structure of a conjugated chain semiconductor caused by the localization of the double bond. The energy $\hbar\omega$ is the energy of the transitions of the side chains.

superconducting state is possible if V is negative and

$$1<\sum_{\epsilon}\frac{|V|}{2[\epsilon^2+(\delta/2)^2]^{1/2}}, \qquad (3.1)$$

where the sum is limited by $|\epsilon'|<\hbar\omega$.

Transforming to an integral, we require

$$\frac{1}{N(0)|V|}<\int_0^x\frac{d\epsilon}{[\epsilon^2+(\delta/2)^2]^{1/2}}, \qquad (3.2)$$

where $x^2=(\hbar\omega)^2-(\delta/2)^2$ and $N(0)$ is defined above prior to the introduction of the gap.

$$\frac{1}{N(0)|V|}<\ln\left[\frac{[(\hbar\omega)^2-(\delta/2)^2]^{1/2}+\hbar\omega}{\delta/2}\right]. \qquad (3.3)$$

If $(\delta/2)^2\ll(\hbar\omega)^2$ then the criterion for obtaining the superconducting state is that

$$\delta<4\hbar\omega\exp\left[\frac{-1}{N(0)|V|}\right]. \qquad (3.4)$$

If this criterion can be satisfied, then the transition temperature can be obtained from the expression

$$\frac{1}{N(0)|V|}=\int_0^x\frac{\tanh\frac{1}{2}\beta_c\epsilon'}{\epsilon'}d\epsilon, \qquad (3.5)$$

which is the same expression for determining the temperature at which the energy gap of a superconductor becomes equal to $\delta/2$. The form of this expression is given graphically in the BCS paper. We only need note that T_c rapidly approaches the transition temperature for $\delta=0$ as $|V|N(0)$ is increased beyond that necessary to satisfy the criterion. So that if the criterion can be satisfied, then the transition temperature generally will be of the same order as the transition temperature of the metallic superconductor. We see then that a gap in the band structure does not necessarily exclude the superconducting state. Incidentally, this conclusion does

[5] Where one does not have alternating double and single bonds, but a double bond separated from the next by two single bonds, the band theory does not appear to work and one finds that the chain is an insulator instead of a metal. This empirical fact does not appear to have been explained, but seems to be a consequence of the same considerations of the metallic and insulating states discussed by N. F. Mott [Phil. Mag. 6, 287 (1961)], and W. Kohn [Phys. Rev. 133, A171 (1964)]. As this question is not quite settled we have limited ourselves to a conjugated chain which is known to behave as one would expect from the band theory.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

not violate Yang's[6] statement that superconductivity or off-diagonal-long-range order (ODLRO) cannot occur in an insulator because his definition of an insulator is one in which there are no available empty states. Our point is that these states are, in fact, available in our model.

It is perhaps appropriate to note that a gap in the density of states such as that considered above would make the average side-chain induced interaction more strongly attractive because for interband transitions $(\epsilon_{k+Q}-\epsilon_k)$ would be at least δ and as can be seen from (2.6), the attractive term would thus become larger provided that $\delta<\hbar\omega$. If one takes the variation of $\langle V(Q)\rangle_{\rm av}$ with Q into account, the details of the above inequality are, of course, changed, but the general feature that the existence of superconductivity is determined by some such criterion remains.

IV. PARTICULAR EXAMPLE

As a particular example of a molecule of the type considered in Sec. II, we will consider in detail the molecule illustrated in Fig. 3. The spine is a conjugated chain of alternating double and single bonds. To this is attached a series of side-chain molecules as shown. Because of the great thickness of the benzene rings in the side chains, compared to the carbon-carbon spacing in the spine, it is not possible to attach a side chain to every carbon atom on the spine. This can be seen more clearly in Fig. 4 where the molecule is drawn to scale using the known values for the van der Waals radii of the constituent atoms. The changed periodicity of the new structure necessitates a slight modification of the wave function used in (2.1) and (2.2) above, but the modification is an obvious one which we will handle later. The side-chain molecule is part of a well-known dye molecule used for sensitizing photographic plates in the red, a diethyl-cyanine iodide, and it has been chosen because its absorption spectrum is known, and its ground and excited states are well understood.[7]

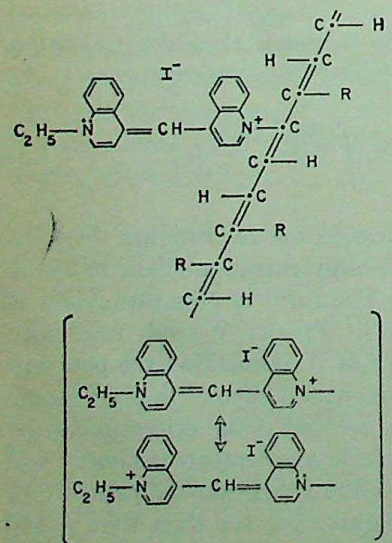


FIG. 3. Chemical structure of the proposed superconducting organic polymer. At each point R on the spine a similar side chain to the one shown is attached. These side chains are resonating hybrids of the two extreme structures shown in the inset. The positive charge resonates between the two nitrogen sites as illustrated.

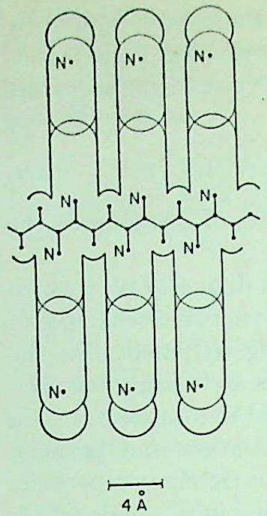


FIG. 4. Approximate scale drawing of the proposed superconducting organic polymer. The plane of the benzene rings in the side chains are oriented at right angles to the spine. The two nitrogen sites on each side chain are indicated, but the iodine site has been omitted for the sake of clarity.

Before estimating the matrix elements of the interaction of the side chains with the electrons in the spine, it is useful to consider first what effect screening will have upon the Coulomb interaction.

Screening

The Coulomb potential at a distance r from a charge which is placed in a conducting medium is screened by the rearrangement of the charges of the medium. The potential is then given approximately by $(e/r)e^{-\lambda r}$. The screening length $1/\lambda$ in a free-electron gas can be estimated by using the Thomas-Fermi method as shown by Mott and Jones.[4] From this one finds that $\lambda^2=(4me^2/\hbar^2)(3N_0/\pi)^{1/3}$ where N_0 is the number of electrons per unit volume. For the carbon atoms on the spine, one valence electron per atom is relatively free, and from the known size of the atom one can estimate N_0 and thus λ. We find that $1/\lambda$ is approximately 0.5 Å. This value varies extremely slowly with the number of valence electrons so that our choice of one free valence electron is not critical. We note then that the Coulomb interaction is screened out in an extremely short distance.

It should be noted too that screening can occur only where the charges of the medium are free to move, as for example, within an atom or along the conjugated series of atoms of the spine. Where the charges are not free to move indefinitely, then only a limited displacement of the charges occurs which merely modifies the Coulomb interaction by the dielectric constant of the medium, but does not screen it. These two cases must be born in mind in considering the Coulomb interaction and the side-chain interaction. Because the electrons are free to move in the spine one must use the screened Coulomb interaction for computing the Coulomb repulsion between the electrons in the spine. However,

[6] C. N. Yang, Rev. Mod. Phys. 34, 694 (1962).

[7] K. Mees, *The Theory of the Photographic Process* (The Macmillan Company, New York, 1942), p. 987.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the interaction between the electrons in the spine and the charges on the side chain is only partially screened because the side chain is insulated from the spine and a free movement of charge between the two cannot occur, i.e., the conjugation of the side chain does not extend to the spine. Here one must use the Coulomb interaction modified by the dielectric constant of the medium together with some screening due to the induced movement of charge in the spine.

Coulomb Repulsion

For steric reasons, it is not possible to attach a side chain to each carbon atom of the spine. Consequently, the Hamiltonian of the spine is not invariant under a displacement from one atom to the next as assumed in our earlier discussion. In our example, the unit cell is repeated only after four carbon atoms and thus, we are dealing here with a lattice with a basis of four atoms. The wave function of the spine is then

$$\phi_{km}=\frac{1}{(G)^{1/2}}\sum_{i}^{G}e^{ikR_i}\sum_{j=1}^{4}\alpha_{kj}U_m(r-r_{j,i}). \qquad (4.1)$$

R_i is now the position of the lattice point measured along the zig-zag line joining the carbon atoms and $r_{j,i}$ the position of the jth atom in the unit cell measured with respect to R_i. α_{kj} is a phase factor which one would expect would be very nearly $(1/4^{1/2})e^{ik(r_{j,i})}$, which is the value it would have if the Hamiltonian was perfectly invariant under a C−C displacement. The number of unit cells, G is now a quarter of the number of carbon atoms in the spine.

Let us now calculate the Coulomb repulsion between electrons in states given by (4.1):

$$V(Q)_{\text{Coulomb}}=\int\int\phi_{k-Q}^{*}(r_1)\phi_{k'+Q}^{*}(r_2)V(r_{12})\times\phi_{k'}(r_2)\phi_k(r_1)d^3r_1d^3r_2. \qquad (4.2)$$

Using (4.1) and the fact that the Coulomb interaction is screened from all except immediate neighboring ions, we obtain in the long-wave limit, i.e., $Q\approx 0$

$$V(0)_{\text{Coulomb}}=\frac{1}{4G}\left[\int|U(r_1-r_j)|^2V(r_{12})|U(r_2-r_j)|^2d^3r_{12}+\sum_{k=j\pm1}\int|U(r_1-r_j)|^2V(r_{12})|U(r_2-r_k)|^2d^3r_{12}\right]. \qquad (4.3)$$

The second term should be much smaller than the first because $V(r)$ is heavily screened. To a first approximation we shall ignore it compared to the first. A reasonable exact value of the dominant term could be obtained by using the known form of the sp carbon orbital and evaluating the integral, however, one can obtain a reasonable estimate of the approximate magnitude of the term by considering the electron density in the orbital as constant and occupying a volume of about half the volume of the atom. The integral is easily done and gives $6e^2/\lambda^2R^3$ where R is the van der Waals radius of the carbon atom. Using the value obtained above for λ and a van der Waals radius of 1.5 Å[8] we obtain a value of 6 eV for the integral. This seems to be a reasonable estimate for one might expect it to be comparable to the energy necessary to add one additional electron to a carbon atom. The energy necessary to form the C^- ion is known as the electron affinity A and is related to the ionization energy I and the electronegativity x of the atom by the relation[8] $(I+A)/5.4=x$ where I and A are expressed in electron volts. The ionization energy of carbon is 11.3 eV and x is 2.5 giving an electron affinity for carbon of 2.2 eV. This is of the same order of magnitude estimated above. Taking the larger value of 6 eV to be safe we find

$$V(0)_{\text{Coulomb}}\approx 1.5\text{ eV}/G. \qquad (4.4)$$

For larger values of Q the first term in (4.3) remains unchanged in the tight-binding approximation while the second term which we have neglected above is reduced by an additional factor of $\cos Qa$ where a is the carbon-carbon spacing. Over the whole range of Q, then we can take the Coulomb repulsion to be of the order of that given by (4.4)

Side-Chain Interaction

The wave function of the side chain must be modified in the same way as the wave function for the electrons of the spine. It is now

$$X_{qn}=\frac{1}{G^{1/2}}\sum_{i=1}^{G}e^{iqR_i}\sum_{j=1}^{4}\beta_{q,j}\psi_n(r-r_{j,i},\cdots). \qquad (4.5)$$

At the sites, j, where side chains are attached β is approximately $e^{iqr_{ji}}/2^{1/2}$, and zero where there are no side chains.

The side chain we have chosen is a resonating hybrid of the two extreme structures shown in the inset to Fig. 3. If the wave functions of these two extreme structures are ψ^+ and ψ^-, respectively, then the ground state, ψ_0 is $(1/\sqrt{2})(\psi_++\psi_-)$ and the excited state of interest to us, ψ_1 is the orthogonal hybrid structure $(1/\sqrt{2})(\psi_+-\psi_-)$. In the matrix element (2.4) we require $\psi_1^*\psi_0$ which is simply $\frac{1}{2}[|\psi_+|^2-|\psi_-|^2]$.

Using the new wave functions for the spine (4.1) and (4.5), Eq. (2.4) reduces to

$$V(Q)_{\text{side chains}}=\int\int\sum_{j=1}^{4}\sum_{n=-1}^{+1}\alpha_{k+Q,j}^{*}\alpha_{kj}|U_m(r-r_{ji})|^2\times V(r)e^{iQn4a\frac{1}{2}}\sum_{l=1}^{4}\beta_{q+Q,l}^{*}\beta_{q,l}\times\{|\psi_+(r_{l,i+n})|^2-|\psi_-(r_{l,i+n})|^2\}, \qquad (4.6)$$

[8] L. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1960), pp. 257, 95.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

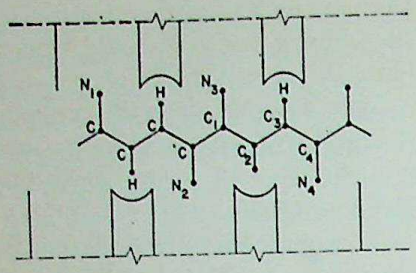


FIG. 5. Enlarged view of atoms of the spine of the proposed polymer.

where we have confined ourselves to a single atomic level m. The different k levels then describe the bands of the type discussed in Sec. III. We recall that j and l refer to each of the four atomic sites within the unit cell; and n has values 0 and ± 1 which refer to the same unit cell, i and the two adjacent cells, respectively.

In the superconducting state two types of matrix elements are important; those which describe scattering between states close to the Fermi surface with momentum transfer $Q \approx 0$ and those for scattering from one side of the Fermi surface to the other. The latter scattering involves a $Q \approx \pi/a$ because there is one electron per atom in the band originating from the fourth sp_3 orbital of the carbon atom. We shall estimate the side-chain interaction matrix element in the neighborhood of these values of the momentum transfer.

Firstly, the interaction between the spine and the iodine ion of the side chain cancels in the matrix element (4.6) because the ion is located at the same point in the two extreme structures of the hybrid. Secondly, in the ψ_- structure the positive charge is located on the nitrogen site which is remote from the spine, consequently, the interaction with this structure is weak. The only important contribution to the matrix element then, is that which is due to the positive charge on the nitrogen close to the spine in the ψ_+ structure, and the electron on each of the sites r_j.

In Fig. 5 we show an enlarged drawing of the spine and part of the side chains. In view of our earlier discussion of screening, it would be consistent to treat the interaction between the positive charge on nitrogen site N_3 (see Fig. 5) and carbon site 1 as virtually unscreened; the interaction with sites 2 and 3 as half screened and sites beyond this as completely screened. Our choice is based on the picture that the migration of negative charge from the vicinity of C_2 and C_3 towards C_1 would partially screen the field produced by the positive charge on N_3 so that at C_4 the interaction would be effectively reduced to zero. We arrive at the figure of $\frac{1}{2}$ for the screening of sites 2 and 3 because half of the atomic orbital adjacent to the nitrogen site on these sites is practically unscreened, while the opposite side is quite strongly screened. There is very little material between the nitrogen site and each of these three carbon sites, so that it seems reasonable to leave the Coulomb interaction with the dielectric constant that of free space. In Table I we tabulate the distances between the nitrogen site and each carbon site. We include also the matrix element for each site, j, in Eq. (4.5) computed for $Q=0$ and $Q=\pi/a$ using the approximate form of the α's and β's given after Eq. (4.1) and (4.5) and the unmodified Coulomb interaction limited in the manner described above.

TABLE I. Carbon-nitrogen distances and matrix elements.

Carbon site	Distance to N_3 (Å)	Matrix element $V(Q)$ (eV) $Q=0$	$Q=\pi/a$
1	1.5	0.76	0.41
2	2.5	0.47	0.11
3	3.0	0.47	0.11
4	(4.5)	0.76	0.41
▮		Total 2.46	1.04

It is reasonable to neglect the interaction with the positive charge on the remote nitrogen site on the ψ_- structure because the distance to the spine is about 14 Å and the interaction is further reduced by the dielectric constant of the material of the side chain. This dielectric constant must be similar to that of benzene which is about 2.2. This gives a total matrix element of 0.1 eV for $Q=0$ which we can neglect compared to the total computed in Table I considering the approximations we have made.

In order to calculate the side-chain induced electron-electron interaction, we must know the energy $\hbar\omega$ for the transitions of the side chains. For an isolated side-chain molecule of 1,1′-diethyl-4,4′-cyanine iodide the absorption maximum for the transition we are considering occurs at 600 mμ giving a value of 2 eV for $\hbar\omega$.[7]

In the polymer the side chains will interact with one another and change the frequency of oscillation to some extent, but as the molecules are quite well separated and the charges are quite well screened from one another, let us take the frequency to be about 2 eV nevertheless. Then the side-chain induced attraction $V \approx \langle 2V(Q)^2 \rangle_{\text{av}}/G\hbar\omega$ will be approximately -3.5 eV/G where we have taken the mean of the square of the interaction for $Q=0$ and $Q=\pi/a$ as given in Table I. This is greater than the Coulomb repulsion of 1.5 eV/G estimated in (4.4) so that the net interaction is an attractive one,

$$V \approx -2 \text{ eV}/G. \tag{4.7}$$

The reason this is so strongly attractive is that we have seen to it that the nitrogen sites lie close to the spine so that the matrix element (4.6) is large and at the same time have chosen a side chain with a fairly low-frequency transition so that $\hbar\omega$ is small.

Some idea of the superconducting transition temperature now can be obtained by estimating the density of states for the electrons in the spine, $N(0)$. This can be done in the following way. The spine itself is very similar to a conjugated polyene chain $\{CH=CH-\}_n$ except that the side chains replace certain of the hydrogens so that one can crudely estimate the density

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of states in the spine from a knowledge of it in a polyene chain. The benzene ring is essentially a triene {CH=CH−} tied back on itself in the form of a loop and can be described in terms of a simple band picture.[9,10] The first absorption at 250 mμ is believed to correspond to a transition of a π-electron from a $m=1$ to $m=2$ state which is approximately half the total width of the band. The total width of the band for benzene is thus ≈10 eV and should be approximately the same for the polyene chain. If we assume the ϵ versus k curve is parabolic up to the halfway point in the band (up to 5 eV), then the density of states of one spin at the Fermi surface is approximately $\frac{1}{5}G$ states/eV. This is probably a reasonable estimate of the density of states as it corresponds to an effective mass for the electrons in the spine of 0.7 the electron mass.

If there was no bond localization in the spine, then we could use Eq. (2.9) to estimate the superconducting transition temperature using (4.3) and the above density of states. One obtains a temperature ≈2200°K in this case! This extremely high transition temperature can be understood when it is realized that in the chosen structure it is an electronic oscillation which provides the coupling between the electrons rather than the oscillation of the nuclei as in a conventional superconductor. The simple argument of the isotope effect that the transition temperature for a phonon-coupled superconductor is proportional to $1/M^{1/2}$, where M is the isotopic mass of the nuclei indicates that for an electron-coupled superconductor the transition temperature should be a factor of $(M/m)^{1/2}$ (i.e., ≈300) times larger. This is, perhaps, too glib an answer for it is necessary to choose the over-all structure so as to obtain a sufficiently strong coupling matrix element (2.6). Our particular model illustrates this in detail.

If there is considerable bond localization in the spine, then our inequality (3.4) shows that the superconducting state can still occur if the semiconductor gap δ is somewhat less than 0.67 eV. If this is satisfied, the transition temperature in this case should still be several hundred °K.

For transition temperatures as high as this the coherence energy of the superconducting state becomes comparable to the chemical binding energy. This energy is approximately

$$W_0=-2N(0)(\hbar\omega)^2\exp[-2/(N(0)|V|)]. \quad (4.8)$$

In our example, the coherence energy is about 0.1 eV per unit cell of the chain. This is not very large, however, if one synthesized a polymer in which the density of states is large but $|V|$ is small so as to obtain the same transition temperature, then the coherence energy would become quite large. A coherence energy of as much as 1 eV per unit cell of the chain appears possible in such a polymer. As this energy is comparable to the resonance energy of the benzene ring, one should expect a considerable stabilization of the polymer on this account. It is interesting to note, too, that the destruction of superconductivity at one point in the chain raises the energy by the coherence energy per unit length times the coherence length ζ_0. The coherence length ζ_0 for these molecules should be about 30 Å as it is inversely proportional to T_c and $\zeta_0\approx10^4$ Å for conventional superconductors. Consequently, it would require a large amount of energy to destroy the superconductivity locally.

V. DISCUSSION

We believe that while the estimates for the various matrix elements in the above example are crude, they are not unrealistic. This forces upon us the remarkable conclusion that superconductivity could and should occur in structures such as this even at room temperatures. There are many other possible structures similar to the one shown involving a semiconducting chain for the spine and a dye-like molecule for the side chain which would also be superconducting. It is unlikely that our particular choice described above would be the easiest to synthesize or have the optimum superconducting properties, but it illustrates the possibility in a detailed manner.

In these molecules we should expect the usual electrical properties of a metallic superconductor, however, in order to observe such effects contact would have to be made to the ends of the spine. This could be a difficult problem to solve, but may be possible by cross-linking the spines so as to form a three-dimensional net of the filamentary molecules. Because of the large transition temperature, one would expect the critical field for the destruction of superconductivity to be very high compared to that of conventional superconductors. The highly divided filamentary structure of a bulk sample of the polymer should mask any appreciable Meissner effect. Perhaps the most interesting feature of these molecules, however, lies in the phase correlation of the electron pairs throughout the molecule. This phase correlation should impose certain restraints upon the ability of the molecule to react chemically with other such molecules. The reason is that in order to form a covalent bond, the electrons must interfere constructively in the region of positive potential. Consequently, the relative phase of the electrons forming the bond are important. Such an effect has been discussed briefly by Ambegaokar and Baratoff[11] in regard to tunneling between conventional superconductors in the Josephson effect. The superconducting state is unique in that this long-range phase correlation, "off-diagonal-long-range-order" (ODLRO)[6] distinguishes it from the normal or insulating states. This we have shown can occur even

[9] J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill Book Company, Inc., New York, 1963), p. 234.

[10] C. R. Noller, *Chemistry of Organic Compounds* (W. B. Saunders Co., Philadelphia, 1960), 2nd ed., p. 665.

[11] V. Ambegaokar and A. Baratoff, Phys. Rev. Letters **10**, 486 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in our structure which is essentially a one-dimensional chain, which is in striking contrast to a classical one-dimensional interacting chain such as the Ising chain which cannot exhibit long-range order.[12] Because of this ODLRO, which is a property of the molecule as a whole and the above chemical properties which are related to it, these molecules will have the property of reacting as a single entity which is precisely what London[1] was seeking to understand in regard to biologically important macromolecules.

In regard to the possible biological significance of our results, it is appropriate to mention a theorem which was established by Wigner[13] on the probability of a quantum mechanical system reproducing itself. He succeeded in showing that under two reasonable assumptions, this probability is essentially zero. The relevant assumption is that his "collision matrix" S, which generates the final state from the initial state, is assumed to be a random matrix. This assumption may be violated in the superconducting state because of the singular nature of the pair distribution associated with ODLRO. It would be useful to reexamine Wigner's theorem to see whether a *superconducting* quantum system would be capable of reproducing. The curious chemical selectivity mentioned earlier suggests that this may be the case.

ACKNOWLEDGMENTS

I should like to thank Derek Griffiths for many useful discussions and critical comments during the preparation of this work. I should also like to thank William Fairbank, Bascomb Deaver, and Felix Bloch for many stimulating conversations over the past few years related to this problem.

Note added in proof. It is appropriate to clarify a point in regard to the band structure of our chosen model. The partial charges on the nitrogen sites of the side chains produce a periodically varying potential along the spine with a fundamental period of $4a$. This produces a gap at $k=\pi/4a$ but no gap at $\pi/2a$ where the Fermi surface lies. The reason there is no gap at $\pi/2a$ is that in the tight binding approximation, where one can consider an electron as sampling the potential at all points within a particular orbital before moving on, the effective potential of each orbital may be taken as that of the appropriate carbon nucleus. From a calculation similar to that of Table II, one can easily show that there is then no Fourier component of period $2a$ and therefore no gap at the Fermi surface. The only effect which tends to produce a semiconductor gap (i.e., one at the Fermi surface) is the tendency for the double bond to localize at alternate sites. This was considered in Sec. III, where we showed that for a fixed semiconductor gap, the superconducting state should occur if a particular inequality could be satisfied. The band gap produced by bond alternation, however, is not fixed, but depends upon the amplitude of the periodic distortion; consequently, the actual gap and distortion must be determined in a self-consistent manner. In this case our earlier arguments do not apply and whether the superconducting state or the semiconductor state occurs depends upon which has the lower energy. Longuet-Higgins and Salem [Proc. Roy. Soc. (London) **A251**, 172 (1959)] have calculated the stabilization energy for a polyene with alternation of bond lengths and obtain an energy of 0.019 kcal/mole per bond ($\approx$0.001 eV/bond) for the semiconductor state. We have calculated it for the superconducting state [Eq. (4.8)] and obtain ($\approx$0.025 eV/bond) so that, in our model, the superconducting state should be favored.

Finally, our calculation has shown that a phase transition from the normal to the superconducting state should occur even in our one-dimensional system. This is unusual [see, for example, L. van Hove, Physica **16**, 137 (1950)] and one may question whether our result follows because the BCS theory, upon which it is based, is not sufficiently exact. We have some reason for believing that our conclusions are valid nevertheless, but this point requires further investigation.

[12] J. Ashkin and W. E. Lamb, Jr., Phys. Rev. 64, 159 (1943).
[13] E. P. Wigner, in *The Logic of Personal Knowledge* (Polanyi Festschrift) (Routledge and Kegan Paul, London, 1961), p. 231.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cavitation in Liquid Helium*

R. D. Finch, R. Kagiwada, M. Barmatz, and I. Rudnick
University of California, Los Angeles, California

(Received 27 January 1964)

Ultrasonic cavitation has been induced in liquid helium over the temperature range 1.2 to 2.3°K, using a pair of identical transducers. The transducers were calibrated using a reciprocity technique and the cavitation threshold was determined at 90 kc/sec. It was found that this threshold has a sharp peak at the λ point, but is at all temperatures quite low, with an approximate range of 0.001–0.01 atm. The significance of these results is discussed.

INTRODUCTION

ESTIMATES of the tensile strengths of liquids based on the assumption that it is possible to rupture molecular bonds, are considerably higher than measured values. For example, in the case of water at room temperatures such predictions of tensile strength vary from 500 to 10 000 atm while empirical values range from a fraction of an atmosphere to 280 atm.[1,2] To explain these discrepancies it has been proposed that true rupture does not occur in practice but that the applied tension is relieved by the sudden growth of relatively large cavities from microscopic nuclei. It is possible to nucleate cavitation in normal liquids by the following principal agencies: (1) minute gas or vapor filled bubbles, mechanically or electrolytically created, (2) inclusions of gas or vapor in cracks on solid surfaces, (3) solid particles unwetted by the liquid, (4) cosmic rays and radioactivity.

The last mentioned agency was first suggested since even degassed distilled liquids have relatively low thresholds.[3] There is now a growing body of phenomena in which cosmic rays and radioactivity nucleate cavitation.[4–10]

It has been pointed out by Beams[11] that the variety of possible nuclei is much restricted in liquid helium since it cannot contain bubbles filled with anything but its own vapor, it is a very efficient wetting agent and due to its low viscosity and density cannot readily carry solid particles in suspension. If cavitation in liquid helium is nucleated, then the most likely agency would be cosmic rays.

When the vapor above liquid helium I is pumped away, evaporative cooling occurs, and bubbles rise through the liquid. On reaching the λ point the bubbling stops and below this temperature the liquid, helium II, has excellent heat transfer properties, causing the disappearance of bubbles.[12] This and other manifestations of superfluidity make it especially interesting for cavitation studies.

The tensile strength of helium II was first measured by Misener and Herbert[13] in the temperature range 1.092–2.147°K using a bellows technique. They concluded that if helium II has a tensile strength it could not exceed 0.3 atm. Beams[11] measured the tensile strength of helium II by a spinning capillary method and by a piston and cylinder method at a temperature of about 1.9°K. He found the negative pressures necessary for rupture were 0.14±0.02 atm with the first method and about one-half of this value with the second method. Since these values are much less than theoretical estimates, Beams concluded that there must have been nuclei present. During preliminary experiments for the design of a liquid-helium bubble chamber Fairbank *et al.*[14] applied negative pressures of about 100 mm to helium II and obtained photographs of bubbles. No tracks were found but there were some "double" bubbles, which Fairbank *et al.* conjectured might have some relationship to quantized vortices.

The uncertainties in these experiments arose from difficulties in the measurement of small negative pressures and in the detection of the onset of cavitation. Relatively low negative pressures can be readily applied and measured using ultrasonic techniques. Further, ultrasonic cavitation gives rise to a characteristic noise which is generated by the collapse of cavities during the positive pressure phase. This noise, when amplified, affords a sensitive means of detecting the onset of cavitation.

APPARATUS

The object of the present experiments was, then, to measure the threshold of ultrasonic cavitation in

* Work supported by the U. S. Office of Naval Research.

[1] B. S. Webster, Ultrasonics 1, 39, (1963).

[2] B. E. Noltingk, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1962), Vol. XI; Acoustics 2, 259 (1962).

[3] W. J. Galloway, J. Acoust. Soc. Am. 26, 849 (1954).

[4] D. Lieberman, Phys. Fluids 2, 466 (1959).

[5] A. L. Hughes, *Proceedings of an International Conference on Instrumentation for High Energy Physics, Berkeley* (Interscience Publishers, New York, 1961), p. 99.

[6] B. Hahn, Nuovo Cimento 22, 650 (1961).

[7] D. Sette and F. Wanderlingh, Phys. Rev. 125, 409 (1962).

[8] D. Messino, D. Sette, and F. Wanderlingh, J. Acoust. Soc. Am. 35, 926 (1963).

[9] B. Hahn and R. N. Peacock, Nuovo Cimento, 28, 334 (1963).

[10] D. Messino, D. Sette, and F. Wanderlingh, J. Acoust. Soc. Am. 35, 1575 (1963).

[11] J. W. Beams, Phys. Rev. 104, 880 (1956).

[12] K. R. Atkins, *Liquid Helium* (Cambridge University Press, Cambridge, England, 1959).

[13] A. D. Misener and G. R. Herbert, Nature 177, 946 (1956).

[14] W. M. Fairbank *et al.*, *Problems of Low Temperature Physics and Thermodynamics* (Pergamon Press, Inc., New York, 1958), p. 45.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

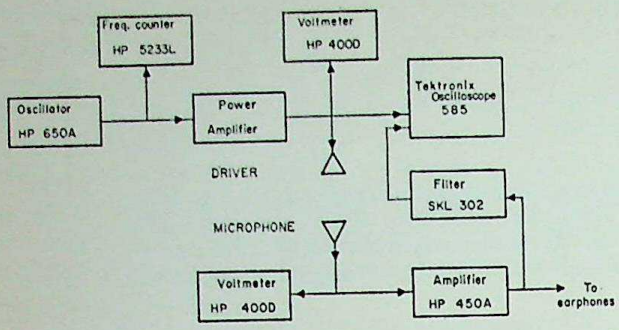


FIG. 1. Block diagram of driver and microphone circuits. With identical PZT-4 transducers as driver and microphone it was possible to deduce acoustic pressures directly from measurements of driver and microphone voltages using a reciprocity technique. The cavitation threshold was determined by the onset of noise from the microphone as heard through earphones and as displayed on an oscilloscope.

liquid helium over a temperature range. To do this two transducers were employed, one to act as a driver and the other as a microphone. A block diagram of the driving and microphone circuits is given in Fig. 1. As will be explained later, it was possible to measure acoustic pressures simply by monitoring the driver and microphone voltages, while the threshold could be detected with cavitation noise, heard in earphones and seen on an oscilloscope trace.

The two transducers were identical disks of Clevite PZT-4 ceramic, $\frac{1}{2}$ in. in radius and in thickness, with electrodes on the flat faces. The resonant frequencies of the transducers were both within 0.2% of 91.15 kc/sec in liquid helium and their capacitances were also the same to within 0.5%. Tests made with a third transducer showed that their sensitivities as microphones were also essentially the same. They were supported by threads (for vibration isolation) from a metal cradle with their opposed faces parallel, there being a separation of some 4.5 cm between them, as seen in Fig. 2. This figure also shows a carbon resistor which when immersed in helium formed one arm of a bridge so that by monitoring the off-balance current an accurate indication of temperature change was possible. The assembly shown, in Fig. 2, was suspended vertically in a standard low-temperature double-Dewar system. The assembly could be viewed through parallel vertical slits in the silvering of the Dewars. Temperatures were determined by measuring vapor pressure using Wallace and Tiernan gauges.

PRELIMINARY OBSERVATIONS OF CAVITATION

When the driving voltage was increased from zero there came a point at which intermittent bursts of noise could be hear in the earphones. This point corresponded to the cavitation threshold and could also be observed as "noise" on an oscilloscope trace of the microphone output after it had been filtered to eliminate the 91-kc/sec driving frequency. With further increase of the driving current the bursts became a steady white noise hiss, characteristic of ultrasonic cavitation, this noise being easily distinguished from that of boiling above the λ point, when the appearance of relatively large rising bubbles could be correlated with "burbling" sounds. With driving currents an order of magnitude or so higher than at threshold the cavitation noise was sufficiently loud to be heard outside the cryostat with an unaided ear. At these power levels the cavities were visible as transient bubbles of about $\frac{1}{2}$–1 mm in diameter, uniformly distributed throughout the volume between the transducers. This was the case whichever transducer was driven. It should be noted, however, that it was not easy to see these bubbles due to the low refractive index of liquid helium, and the use of photography was not practical since the cryostat could not readily accommodate optical systems of any size. With these reservations in mind it may be stated that no "double" bubbles were seen. An examination was made with an RCA 6810-A photomultiplier (cathode sensitivity: 60 μA/lumen, current gain: 12.5×10^6) but no sonoluminescence could be detected at any temperature in the operating range.

MEASUREMENT OF CAVITATION THRESHOLDS

It has been shown by Maclean[15] that using the laws of reciprocity for a four-terminal network it is possible to determine the sensitivity of a reversible transducer using only one other identical transducer. Following

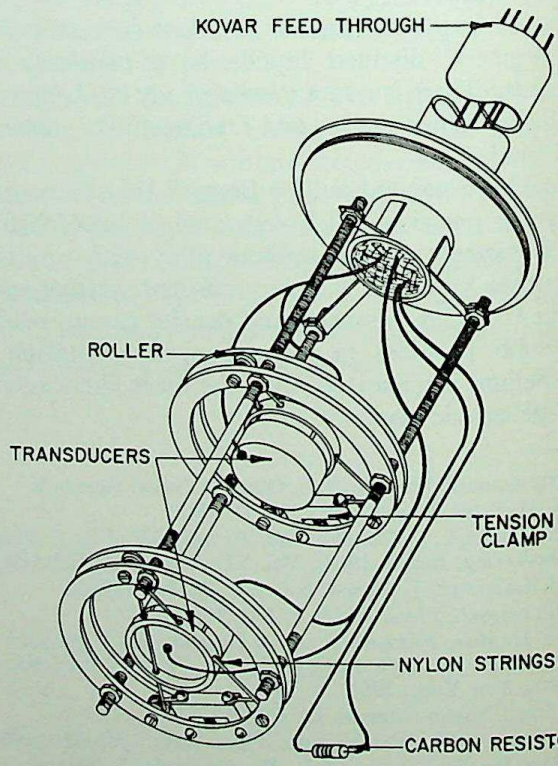


FIG. 2. Driver and microphone suspension. The PZT-4 transducers were 1 in. in diam and $\frac{1}{2}$ in. in thickness. A carbon resistor, used to determine temperature stability, is seen below the lower transducer.

[15] W. R. Maclean, J. Acoust. Soc. Am. 12, (1940).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the principles laid down by Maclean it may be shown that for two identical transducers:

$$M_0=[(V_2/I_1)(1/Q)]^{1/2}, \quad (1)$$

where M_0=sensitivity of a transducer used as a microphone in esu of potential/dyne/cm². V_2 and I_1 are, respectively, the microphone open circuit potential difference and driving current in esu as indicated in Fig. 3. Q is given by

$$Q=P_2/U_1S,$$

where P_2=acoustic pressure (dynes/cm²) at the microphone face, U_1=particle velocity at the driver face (cm/sec) and S=area of a transducer face (cm²).

In the present case a standing wave is set up in the volume between the transducers. However, since the driver has a high impedance the particle velocity U_1 will not differ appreciably from the free field particle velocity U_F. Suppose that due to the nondirectionality of the beam the pressure in a progressive wave is reduced by a factor r in traveling once between the transducers. The pressure at the microphone, P_2, will then be given by

$$P_2=2P_Fr/(1-r^2),$$

where P_F is the free field pressure at the driver. Thus, in this case,

$$Q=(\rho c/S)[2r/(1-r^2)],$$

where ρ and c are the density and sound velocity of liquid helium, respectively. In Fig. 3 C_T represents the capacitance of the transducer and C_g the capacitance of the cable. The capacitances of the cables had essentially the same value (within 0.5% of 524 μF) showing no appreciable change over the operating temperature ranges. Under these circumstances $V_2'I_1'=V_2I_1$ and $I_1'=V_1'Z_D$, where Z_D is the combined impedance of a transducer and its cable. A further test was made to confirm that Z_D did not vary with temperature. Hence, substituting in Eq. (1) for Q and taking values of V_2' and I_1' appropriate to the onset of cavitation then the

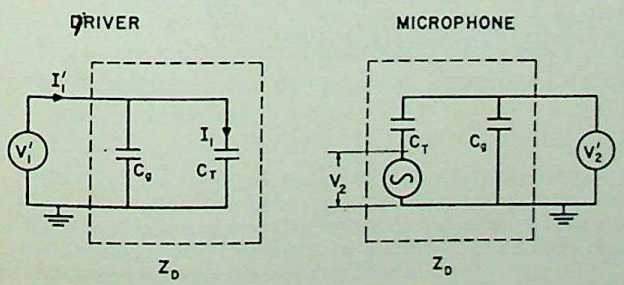


FIG. 3. Four terminal reciprocity network. C_t: capacitance of a transducer. C_g: cable capacitance. C_t and C_g were identical to within ½% for both transducers. Z_d: combined impedance of a transducer and its cable. V_2: generated microphone voltage. V_2': measured microphone voltage. I_1: actual driving current. I_1': measured driving current. V_1' measured driving voltage.

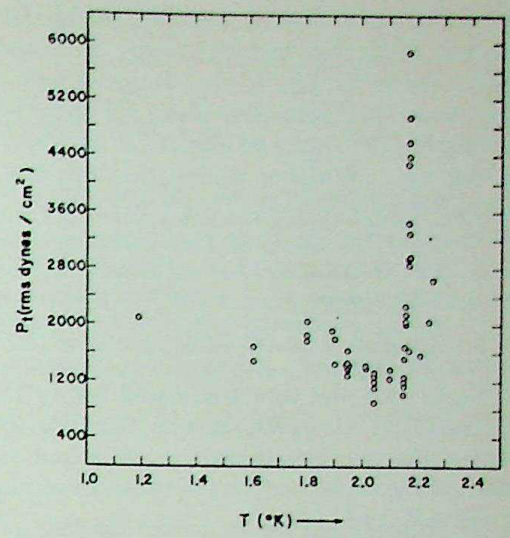


FIG. 4. Cavitation thresholds. The principal feature is a sharp peaking to a maximum value at the λ point (2.172°K). This rise all occurs within 0.01°K below the λ point. At other temperatures the threshold is comparable to the liquid pressure head, which varied smoothly from about 1600 to 1100 dynes/cm² at the driver face during the experiment.

pressure at the microphone, Pm, will be:

$$P_m=\left(\frac{\rho c}{S}\frac{2r}{(1-r^2)}\frac{10^7}{Z_D}\right)^{1/2}(V_1V_2)^{1/2},$$

where V_1 and V_2 are expressed in volts and Z_D in ohms. The pressure at the driver is then the cavitation threshold, P_t, given by

$$P_t=\frac{(1+r^2)}{2r}\left(\frac{\rho c}{S}\frac{2r}{(1-r^2)}\frac{10^7}{Z_D}\right)^{1/2}(V_1V_2)^{1/2}. \quad (2)$$

The ratio r was measured by applying pulses of about 100-μsec width to the driver, without removing the assembly from the helium, displaying the microphone voltage on the oscilloscope and measuring the ratio of successive reflected pulses. This gave r^2, which was found to be 0.69 by averaging a number of measurements, so that $r=0.83$. Then, taking $\rho=0.146$ gm/cm³, $c=2.38\times10^4$ cm/sec, $S=5.07$ cm², and $Z_D=2.75\times10^3$ Ω, from Eq. (2), $P_t=3.72\times10^3(V_1V_2)^{1/2}$. Thus the threshold at any temperature could be deduced directly from two voltage readings.

The procedure adopted was to increase the driving voltage until cavitation noise could just be heard continuously. Then by gradually reducing the driving voltage a point was reached with careful tuning where a burst of noise could not be heard more than once every five seconds. The voltage readings were then made. Certain precautions had to be observed to obtain consistent readings. To begin with, care was taken to ensure that particles of frozen air did not get into the helium. Secondly, it was found that the process of temperature stabilization (achieved by balancing the heat input from the transducer against the pumping

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

rate) had to be truly effective and the threshold measurement not made too hurriedly after stabilization. The results shown in Fig. 4 were obtained only after the temperature had remained stable for some ten to fifteen minutes. The liquid level was 7.2 cm above the face of the top transducer during this run and the bottom transducer was used as the driver.

Some calorimetric measurements were carried out by timing the temperature rise from 1.53 to 1.73°K with the sound pressure maintained just above threshold level, and then with no sound at all. The time intervals in the two cases did not differ by more than about 5%, so that the threshold could not be determined accurately. However, it was found that the total power flux due to self-heating and sound across the surface of the transducer was of the order of magnitude of one mW/cm^2.

DISCUSSION OF RESULTS

In the first place, it is necessary to establish that the phenomenon observed was truly cavitation and not boiling at the surface of the transducer. In order to boil helium II it is necessary to generate a heat current of sufficient magnitude to convert the liquid surrounding the source into helium I after which the boiling occurs. A number of workers have found that this heat current is of the order of 1 W/cm^2 [16,17] and that heat currents of the same order of magnitude are required for nucleate boiling of helium I.[18] In the present experiments the heat current at threshold was of the order of a mW/cm^2. Only when driving the transducer at full power could the heat flux have attained 1 W/cm^2. This is strong evidence that boiling was not present, a conclusion corroborated by a number of facts. For instance the white noise obtained was quite distinct from the sounds of nucleate boiling and typical of ultrasonic cavitation. Again, at high ultrasonic power transient bubbles could be seen uniformly distributed in the space between the transducers, even when driving the upper one. Whereas if the phenomenon were boiling the vapor phase might be expected to occur in the locality of the driving transducer or as stable bubbles rising from it.

A significant aspect of the results is the fact that the threshold is very low at all temperatures. Apart from the λ point the situation is similar to the case of gassy water where the threshold is determined by the static pressure head. However, the methods for estimating the tensile strength of a normal liquid may not be applicable in the case of helium, so that the values of 4 atm and over, as cited by Beams[11] should be treated with some reserve.

The other remarkable feature of the results is the sharp peak at the λ point. Many of the properties of liquid helium show pronounced effects in this vicinity, especially the thermal properties. There is for instance the well-known behavior of the specific heat which rises sharply to very high values at the λ point. One possible tentative explanation of this apparent correlation between the threshold and thermal properties can be given in terms of the Seitz theory of the bubble chamber.[19]

As has been mentioned in the Introduction, of the agents that are known to nucleate cavitation in normal liquids, cosmic rays and radioactivity are the most likely to be operative in liquid helium. That the passage of high-energy particles in liquid helium can give rise to bubbles is evidenced by the successful operation of helium bubble chambers. According to Seitz, the particle produces highly localized energy deposits or "thermal spikes" in the liquid. The thermal spikes then cause the growth of microbubbles whose equilibrium size is partly determined by the thermal properties of the liquid. Possibly it is these microbubbles which nucleate the cavitation event. If the threshold is a function of nucleus size, as in normal liquids, then a relationship might be expected between the threshold and the thermal properties.

ACKNOWLEDGMENT

The authors wish to thank K. A. Shapiro for making available the cryostat used in the experiments.

[16] H. E. Rorschach and F. A. Romberg, *Proceedings of the Fifth International Conference on Low Temperature Physics and Chemistry* (Universiy of Wisconsin Press, Madison, Wisconsin, 1958), p. 35.

[17] F. Haenssler and L. Rinderer, Helv. Phys. Acta 32, 322 (1959).

[18] A. Karagounis, Suppl. Bull. Inst. Int. du Froid, Annexe 2, 195 (1956).

[19] F. Seitz, Phys. Fluids 1, 2 (1958).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Theory of an Optical Maser*

WILLIS E. LAMB, JR.
Yale University, New Haven, Connecticut
(Received 13 January 1964)

A theoretical model for the behavior of an optical maser is presented in which the electromagnetic field is treated classically, and the active medium is made up of thermally moving atoms which acquire nonlinear electric dipole moments under the action of the field according to the laws of quantum mechanics. The corresponding macroscopic electric polarization of the medium acts as a source for an electromagnetic field. The self-consistency requirement that a quasistationary field should be sustained by the induced polarization leads to equations which determine the amplitudes and frequencies of multimode oscillation as functions of the various parameters characterizing the maser. Among the results obtained are: threshold conditions, single-mode output as a function of cavity tuning, frequency pulling and pushing, mode competition phenomena including frequency locking, production of combination tones, and population pulsations. A more approximate discussion of maser action using rate equations is also given in which the concept of "hole burning" plays a role.

1. INTRODUCTION

THIS paper gives a theoretical description of the operation of multimode maser oscillators. The type of approach is particularly suitable for gaseous optical masers of the type suggested by Schawlow and Townes,[1] and first realized experimentally by Javan, Bennett, and Herriott,[2] but the equations should also find use in the description of some features of solid-state optical masers.

2. BASIS FOR CALCULATION

We consider a high-Q multimode cavity in which there is a given *classical* electromagnetic field acting on a material medium which consists of a collection of atoms described by the laws of quantum mechanics. No attempt is made to consider noise due to spontaneous emission and thermal, density, or quantum fluctuations. The high degree of spectral purity observed by Javan and co-workers[3] suggests that these should be good approximations.

The effect of the electromagnetic field on the atoms in the cavity is to produce a macroscopic electric polarization $\mathbf{P}(\mathbf{r},t)$ of the medium. This acts as a source for the electromagnetic field in accordance with Maxwell's equations. The conditions for self-consistency (that the field produced should be equal to the field assumed) determine the amplitudes and frequencies of the possible oscillations. The calculations will include nonlinear effects, so that phenomena of frequency pulling and pushing, mode competition, frequency locking, etc., can be described.

The thermal motion of an atom during its natural decay time may carry it several wavelengths through the standing wave pattern of the electromagnetic field. As a result, the atom "sees" Doppler-shifted optical frequencies which depend on its trajectory. This important circumstance considerably influences the behavior of the Javan-Bennett-Herriott maser. When, however, thermal motion is neglected the equations of the paper can be used in a model calculation for an ideal solid-state optical maser.

We will assume that only two atomic states a and b contribute to the maser action. As a related simplification the vector character (polarization) of the electromagnetic field will be ignored. In order to ensure that our analysis should apply, it would be desirable to have the optical configuration favor one plane of polarization, as with windows of the Brewster's angle type. The more complicated problem of a general state of polarization will be dealt with in another paper.

A cavity of the Fabry-Perot type used by Javan, Bennett, and Herriott has, of course, a continuum of modes because it is not enclosed by reflecting walls. However, it follows from work of Fox and Li[4] that there are discrete sets of quasimodes for which the diffractive leakage from the tube is small. The cavity modes of highest Q are the even symmetric ones whose circular frequencies are given by

$$\Omega_n = \pi n c/L, \tag{1}$$

where c is the velocity of light, L is the distance between the reflecting plates ($L \sim 100$ cm) and n is a large integer, typically of order 2×10^6. Fox and Li have shown that the modes of next highest Q are those possessing odd radial symmetry, which, for typical geometry differ by about 1 Mc/sec from the former modes. Our discussion will be specifically, but not inevitably, aimed at the modes of highest Q.

* This work was supported in part by the U. S. Air Force Office of Scientific Research. The main results of the paper were reported at the Third International Conference on Quantum Electronics, Paris, February, 1963. Lectures on some of the material were given at the 1963 Varenna Summer School.

[1] A. L. Schawlow and C. H. Townes, Phys. Rev. 112, 1940 (1958).

[2] A. Javan, W. R. Bennett and D. R. Herriott, Phys. Rev. Letters 6, 106 (1961).

[3] T. S. Jaseja, A. Javan, and C. H. Townes, Phys. Rev. Letters 10, 165 (1963).

[4] A. G. Fox and T. Li, Bell System Tech. J. 40, 61 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

3. ELECTROMAGNETIC FIELD EQUATIONS

We write Maxwell's equations in mks units as

$$\begin{aligned} &\mathrm{div}\mathbf{D}=0 \quad \mathrm{curl}\mathbf{E}=-\partial\mathbf{B}/\partial t \\ &\mathrm{div}\mathbf{B}=0 \quad \mathrm{curl}\mathbf{H}=\mathbf{J}+\partial\mathbf{D}/\partial t, \end{aligned} \tag{2}$$

where

$$\mathbf{D}=\epsilon_0\mathbf{E}+\mathbf{D}, \quad \mathbf{B}=\mu_0\mathbf{H}, \quad \mathbf{J}=\sigma\mathbf{E}. \tag{2a}$$

To an approximation whose validity will be discussed in another paper, the array of excited atoms may be regarded as a medium with an electrical state described by a macroscopic polarization $\mathbf{P}(\mathbf{r},t)$ (electric dipole moment density). In order to avoid a complicated boundary value problem, it is convenient to assume the presence of a lossy medium with an Ohmic conductivity σ adjusted to give the desired damping of a normal mode. The electric field then obeys a wave equation

$$\mathrm{curl}\,\mathrm{curl}\mathbf{E}+\mu_0\sigma\partial\mathbf{E}/\partial t+\mu_0\epsilon_0\partial^2\mathbf{E}/\partial t^2=-\mu_0\partial^2\mathbf{P}/\partial t^2. \tag{3}$$

In the subsequent calculations the main effect of the space dependence of $\mathbf{E}(x,y,z,t)$ comes from the motion of the excited atoms through the field which leads to amplitude modulation of the fields seen by the atoms. The analysis of Fox and Li for the even symmetric modes indicates that the electric field does not vary rapidly across the tube diameter. Accordingly we take only the axial variation of $\mathbf{E}$ into account. Then curl curl$\mathbf{E}$ is replaced by $-\partial^2E/\partial z^2$, where z is the axial coordinate, and E is the transverse electric field. For the nth normal mode (unnormalized), we have eigenfunctions

$$U_n(z)=\sin K_n z, \tag{4}$$

with wave number

$$K_n=n\pi/L, \tag{5}$$

where n is a large integer.

In the presence of a given polarization $P(z,t)$, quasi-stationary forced oscillations of the electric field can be expanded in normal mode eigenfunctions

$$E(z,t)=\textstyle\sum_n A_n(t)U_n(z), \tag{6}$$

where the amplitudes $A_n(t)$ obey a differential equation of a forced, damped simple harmonic oscillator

$$\frac{d^2A_n}{dt^2}+\left(\frac{\sigma}{\epsilon_0}\right)\frac{dA_n}{dt}+\Omega_n^2A_n=-\left(\frac{1}{\epsilon_0}\right)\frac{d^2P_n(t)}{dt^2}, \tag{7}$$

in which $P_n(t)$ is the space Fourier component of $P(z,t)$

$$P_n(t)=\frac{2}{L}\int_0^L dzP(z,t)\sin K_n z. \tag{8}$$

Since $P_n(t)$ will be very nearly monochromatic at an optical frequency[5] (o.f.) ν, we replace its second time derivative by $-\nu^2P_n$ on the right side of Eq. (7). Adjusting the fictional conductivity σ to give the desired Q_n of the nth mode, we write

$$\sigma=\epsilon_0\nu/Q_n. \tag{9}$$

Then $A_n(t)$ obeys

$$\frac{d^2A_n}{dt^2}+\left(\frac{\nu}{Q_n}\right)\frac{dA_n}{dt}+\Omega_n^2A_n=\left(\frac{\nu^2}{\epsilon_0}\right)P_n. \tag{10}$$

In the typical gaseous optical maser, the separation of the principal modes $\Delta\sim150$ Mc/sec is much larger than the cavity mode band width $\nu/Q\sim1$ Mc/sec. Hence we may hope to neglect time Fourier components of $A_n(t)$ and $P_n(t)$ which are at frequencies far from the cavity resonance frequency Ω_n, and write[6]

$$A_n(t)=E_n(t)\cos(\nu_n t+\varphi_n(t)), \tag{11}$$

and

$$\begin{aligned} P_n(t)=C_n(t)\cos(\nu_n t+\varphi_n(t)) \\ +S_n(t)\sin(\nu_n t+\varphi_n(t)), \end{aligned} \tag{12}$$

where the amplitudes $E_n(t)$ and phases $\varphi_n(t)$, as well as the in-phase and quadrature coefficients $C_n(t)$ and $S_n(t)$ are slowly varying functions of t which, together with the frequencies ν_n, are still to be determined. The expressions (11) and (12) are put into Eq. (10) with only the first time derivatives of $E_n(t)$ and $\varphi_n(t)$ retained. Equating the coefficients of $\cos(\nu_n t+\varphi_n)$ and $\sin(\nu_n t+\varphi_n)$ separately to zero, and further neglecting small terms involving $\nu_n\dot{E}_n/Q_n$, $\dot{\varphi}_n\dot{E}_n$ and $\nu_n\dot{\varphi}_nE_n/Q_n$, and recognizing that $\nu_n+\dot{\varphi}_n$ is very close to Ω_n, we find the self-consistency equations

$$(\nu_n+\dot{\varphi}_n-\Omega_n)E_n=-\tfrac{1}{2}(\nu/\epsilon_0)C_n \tag{13}$$

and

$$\dot{E}_n+\tfrac{1}{2}(\nu/Q_n)E_n=-\tfrac{1}{2}(\nu/\epsilon_0)S_n, \tag{14}$$

which serve to determine the amplitudes, frequencies and phases of the o.f. radiation once the polarization state of the medium is known in terms of the $E_n(t)$.

4. POLARIZATION OF THE MEDIUM

The maser action arises from the establishment of a negative temperature distribution for the two excited states a and b of the atoms constituting the medium as shown in Fig. 1. The ground state, far below a and b, is not shown. Consider what happens to an atom which at time t_0 is excited by some process (electron bombardment, collision of the second kind, absorption of resonance radiation, decay from some higher excited

[5] We adopt the convention that all symbols for frequencies should denote circular frequencies. A numerical value, e.g., 150 Mc/sec, however, denotes an ordinary frequency. A decay constant like γ_a which denotes a reciprocal life time $1/\tau_a$ often plays the role of a circular frequency. Numerical values of γ_a will be given as ordinary frequencies.

[6] The representation of an arbitrary function $A_n(t)$ in the form (11) in terms of a variable amplitude $E_n(t)$ and phase $\varphi_n(t)$ is not unique. Despite this, because of the use of the rotating wave approximation it seems possible through Eqs. (13) and (14) to determine both amplitude and phase. [The positive frequency part of (11) is a complex function very closely equal to $A_n^{(+)}(t)=E_n(t)\exp-i(\nu_n t+\varphi_n(t))$ which does have a unique amplitude and phase.]

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

state, etc.) into the upper maser state a. Let the atom be at position $\mathbf{r}_0$ at t_0, and have velocity $\mathbf{v}$. For the present, we neglect collisions, so that at time $t>t_0$ the atom will be at $\mathbf{r}=\mathbf{r}_0+\mathbf{v}(t-t_0)$. If there is an o.f. electric field $\mathbf{E}(\mathbf{r},t)$ in the cavity, the atom sees a time-dependent field $E(\mathbf{r}_0+\mathbf{v}t-\mathbf{v}t_0, t)$ for $t>t_0$. Associated with this field is a time-dependent perturbation energy whose matrix element is

$$\hbar V(t)=-\wp E(\mathbf{r}_0+\mathbf{v}t-\mathbf{v}t_0, t), \tag{15}$$

where $\wp$ (assumed real) is the matrix element for the electric dipole moment of the atom between states a and b. The perturbation causes the atomic wave function to become a time-dependent linear combination $a(t)\psi_a+b(t)\psi_b$. The quantum-mechanical average value of the electric dipole operator for the atom is $(a^*b+ab^*)\wp$.

To follow the time-dependent wave function (in the subspace of ψ_a and ψ_b), we start from the equations of time-dependent perturbation theory

$$\begin{aligned} i\dot{a}&=W_a a+V(t)b-\tfrac{1}{2}i\gamma_a a,\\ i\dot{b}&=W_b b+V(t)a-\tfrac{1}{2}i\gamma_b b, \end{aligned} \tag{16}$$

in which the radiative decay of states a and b is described by phenomenological terms containing the decay constants γ_a and γ_b for the two states. Here $\hbar W_a$ and $\hbar W_b$ are the unperturbed energies of states a and b, and the matrix element of the perturbation $V(t)$ is given by Eq. (15).

If the motion of the atom were neglected, and if the maser were working in a single cavity mode, $V(t)$ would be monochromatic, and the rotating wave approximation would allow the Eqs. (16) to be integrated exactly. Even so, there are great algebraic simplifications to be gained by going over to a density matrix description[7] of an ensemble of atoms consisting of all those of a given category which are produced during all times $t_0<t$. A theory of maser action in this case has already been given[8] which is valid when the signals are strong enough to fully saturate the transition $a\leftrightarrow b$. For multimode operation, such an exact solution can no longer be obtained. However, the simpler theory can help with the interpretation of our rather complicated equations, and it will be discussed in Secs. 16–20.

When atomic motion through the electromagnetic field is taken into account an atom does not see a monochromatic perturbation even in single-mode operation. The equations can only be solved in a perturbation expansion of the solution in powers of the $E_n(t)$. It is still advantageous to use the density matrix method, considering first only those atoms characterized by

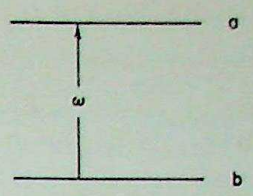


FIG. 1. Two excited energy levels a and b between which the maser action takes place. The levels have a resonance transition frequency $\omega>0$, and are given phenomenological decay constants γ_a and γ_b. The excitation of the states is described by the functions $\lambda_\alpha(r_0,t_0,v)$ which are introduced in Eq. (22).

a, $\mathbf{r}_0$, t_0, $\mathbf{v}$. The density matrix

$$\begin{aligned} \rho(a,\mathbf{r}_0,t_0,\mathbf{v},t)&=\begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}\\ &=\begin{pmatrix} \rho_{aa} & \rho_{ab} \\ \rho_{ba} & \rho_{bb} \end{pmatrix} \end{aligned} \tag{17}$$

obeys an equation of motion

$$\dot{\rho}=-i[\mathcal{H},\rho]-\tfrac{1}{2}(\Gamma\rho+\rho\Gamma), \tag{18}$$

where Γ is the diagonal matrix

$$\Gamma=\begin{pmatrix} \gamma_a & 0 \\ 0 & \gamma_b \end{pmatrix} \tag{19}$$

and the Hamiltonian matrix $\mathcal{H}$ is

$$\mathcal{H}=\begin{pmatrix} W_a & V(t) \\ V(t) & W_b \end{pmatrix}, \tag{20}$$

with $V(t)$, as given by Eq. (15), having a complicated time dependence because of atomic motion. A solution of Eq. (18) which satisfies the initial conditions

$$\rho(a,\mathbf{r}_0,t_0,\mathbf{v},t_0)=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{21}$$

is required. The average electric dipole moment corresponding to this density matrix ρ is $\wp(\rho_{ab}+\rho_{ba})$.

To obtain the macroscopic polarization $P(\mathbf{r},t)$ we have to combine the contributions of all atoms which arrive at $\mathbf{r}$ at time t, no matter when or where they were excited to state a, and also a similar contribution from atoms excited initially to state b. Let $\lambda_\alpha(\mathbf{r}_0,t_0,\mathbf{v})$ be the number of atoms excited to state $\alpha=a, b$ per unit time per unit volume. We have

$$\begin{aligned} P(\mathbf{r},t)=\wp\sum_{\alpha=a,b}\int_{-\infty}^{t}dt_0\int d\mathbf{r}_0\int d\mathbf{v}\lambda_\alpha(\mathbf{r}_0,t_0,\mathbf{v})\\ \times[\rho_{ab}(\alpha,\mathbf{r}_0,t_0,\mathbf{v},t)+\rho_{ba}(\alpha,\mathbf{r}_0,t_0,\mathbf{v},t)]\\ \times\delta(\mathbf{r}-\mathbf{r}_0-\mathbf{v}(t-t_0)). \end{aligned} \tag{22}$$

In practice, $\lambda_\alpha(\mathbf{r}_0,t_0,\mathbf{v})$ will be a slowly varying function

[7] W. E. Lamb, Jr. and T. M. Sanders, Jr., Phys. Rev. 119, 1901 (1960), especially pp. 1902–1903; L. R. Wilcox and W. E. Lamb, Jr., *ibid.* 119, 1915 (1960), especially p. 1928.

[8] W. E. Lamb, Jr., *Quantum Mechanical Amplifiers, in Lectures in Theoretical Physics*, edited by W. E. Brittin and B. W. Downs (Interscience Publishers, Inc., New York, 1960), Vol. II, especially pp. 472–476.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of $\mathbf{r}_0$ so that it can be replaced by $\lambda_\alpha(\mathbf{r},t_0,\mathbf{v})$. After integration over $\mathbf{r}_0$

$$P(\mathbf{r},t)=\wp\sum_{\alpha=a,b}\int_{-\infty}^{t}dt_0\int d\mathbf{v} \tag{23}$$
$$\times\lambda_\alpha(\mathbf{r},t_0,\mathbf{v})[\rho_{ab}(\alpha,\ \mathbf{r}-\mathbf{v}(t-t_0),\ t_0,\ \mathbf{v},\ t)+\text{conj.}].$$

Similarly, we will have use for a density matrix describing an ensemble of atoms which arrive at $\mathbf{r}$ with velocity $\mathbf{v}$ at time t regardless of their place $\mathbf{r}_0$, time t_0 or state $\alpha=a,\ b$ of excitation. This will be denoted by

$$\rho(\mathbf{r},\mathbf{v},t)=\sum_{\alpha=a,b}\int_{-\infty}^{t}dt_0\int d\mathbf{r}_0\lambda_\alpha(\mathbf{r}_0,t_0,\mathbf{v})$$
$$\times\rho(\alpha,\mathbf{r}_0,t_0,\mathbf{v},t)\delta(\mathbf{r}-\mathbf{r}_0-\mathbf{v}t+\mathbf{v}t_0). \tag{24}$$

The density matrix resulting from (24) by integration over all velocities will be denoted by $\rho(\mathbf{r},t)$.

5. INTEGRATION OF THE EQUATIONS OF MOTION

The matrix equation of motion for the density matrix $\rho(\alpha,\mathbf{r}_0,t_0,\mathbf{v},t)$ has components

$$\dot\rho_{ab}=-i\omega\rho_{ab}-\gamma_{ab}\rho_{ab}+iV(t)(\rho_{aa}-\rho_{bb}),$$
$$\dot\rho_{aa}=-\gamma_a\rho_{aa}+iV(t)(\rho_{ab}-\rho_{ba}), \tag{25}$$
$$\dot\rho_{bb}=-\gamma_b\rho_{bb}-iV(t)(\rho_{ab}-\rho_{ba}),$$
$$\rho_{ba}=\rho_{ab}^*, \tag{26}$$

where

$$\gamma_{ab}=\tfrac{1}{2}(\gamma_a+\gamma_b) \tag{26a}$$

and

$$\omega=W_a-W_b>0. \tag{27}$$

We consider first the case of excitation to the upper maser state a. At $t=t_0$, $\rho_{aa}=1$ and $\rho_{bb}=\rho_{ab}=\rho_{ba}=0$. The solution to any desired order in the perturbation $V(t)$ can be obtained by iteration. There are contributions to $\rho_{ab}=\rho_{ba}^*$ in first and third order, to ρ_{bb} in second order, and to ρ_{aa} in zeroth and second order. Thus, in zeroth order,

$$\rho_{aa}^{(0)}(a,\mathbf{r}_0,t_0,\mathbf{v},t)=\exp-\gamma_a(t-t_0) \tag{28}$$

and the first-order contribution to ρ_{ab} is

$$\rho_{ab}^{(1)}(a,\mathbf{r}_0,t_0,\mathbf{v},t)$$
$$=i\int_{t_0}^{t}dt'V(t')\exp[(\gamma_{ab}+i\omega)(t'-t)+\gamma_a(t_0-t')], \tag{29}$$

while the third-order contribution is

$$\rho_{ab}^{(3)}(a,\mathbf{r}_0,t_0,\mathbf{v},t)=i\int_{t_0}^{t}dt'V(t')$$
$$\times[\rho_{aa}^{(2)}(a,\mathbf{r}_0,t_0,\mathbf{v},t')-\rho_{bb}^{(2)}(a,\mathbf{r}_0,t_0,\mathbf{v},t')]$$
$$\times\exp(\gamma_{ab}+i\omega)(t'-t), \tag{30}$$

where

$$\rho_{aa}^{(2)}(a,\mathbf{r}_0,t_0,\mathbf{v},t')$$
$$=-\int_{t_0}^{t'}dt''\int_{t_0}^{t''}dt'''V(t'')V(t''')$$
$$\times\{\exp[\gamma_a(t''-t')+(\gamma_{ab}+i\omega)(t'''-t'')$$
$$+\gamma_a(t_0-t''')]+\text{conj.}\}, \tag{31}$$

and similarly

$$\rho_{bb}^{(2)}(a,\mathbf{r}_0,t_0,\mathbf{v},t')$$
$$=+\int_{t_0}^{t'}dt''\int_{t_0}^{t''}dt'''V(t'')V(t''')$$
$$\times\{\exp[\gamma_b(t''-t')+(\gamma_{ab}+i\omega)(t'''-t'')$$
$$+\gamma_a(t_0-t''')]+\text{conj.}\}. \tag{32}$$

6. FIRST-ORDER THEORY

In order to convert the expression (29) for $\rho_{ab}^{(1)}$ into a macroscopic polarization, we must first calculate

$$\rho_{ab}^{(1)}(a,\mathbf{r},\mathbf{v},t)$$
$$=\int_{-\infty}^{t}dt_0\rho_{ab}^{(1)}(a,\ \mathbf{r}_0=\mathbf{r}-\mathbf{v}t+\mathbf{v}t_0,\ t_0,\ \mathbf{v},\ t) \tag{33}$$

as in Eq. (24).

The perturbation $V(t')$ acting at time t' on the atom specified by $\mathbf{r}_0$, t_0, $\mathbf{v}$ is $-(\wp/\hbar)E(\mathbf{r}_0+\mathbf{v}(t'-t_0),t')$, but for Eq. (33) we require this for an atom characterized by $\mathbf{r}_0=\mathbf{r}-\mathbf{v}t+\mathbf{v}t_0$, $\mathbf{v}$, t_0, for which the effective perturbation

$$V(t')=-(\wp/\hbar)E(\mathbf{r}-\mathbf{v}(t-t'),t') \tag{34}$$

does not depend on t_0. We may then perform the above integration over t_0 if we treat λ_a as a slowly varying function of t_0 and evaluate it at t. For the first-order terms, we have to deal with an expression of the form

$$\int_{-\infty}^{t}dt_0\int_{t_0}^{t}dt'F(t,t')e^{\gamma_a(t_0-t')}.$$

By an interchange[9] of the order of integrations this becomes

$$\int_{-\infty}^{t}dt'\int_{-\infty}^{t'}dt_0F(t,t')e^{\gamma_a(t_0-t')}=(1/\gamma_a)\int_{-\infty}^{t}dt'F(t,t'). \tag{35}$$

Let us assume now that the maser oscillator is running simultaneously in M cavity modes, so that

$$E(\mathbf{r},t)=\sum_{\mu=1}^{M}E_\mu(t)U_\mu(\mathbf{r})\cos(\nu_\mu t+\varphi_\mu(t)), \tag{36}$$

[9] In each double integral an integration is carried over the same triangular area in the t_0, t' plane.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where $E_\mu(t)$ and $\varphi_\mu(t)$ are slowly varying functions of time. We make a rotating wave approximation by keeping only exponential factors like $\exp i(\omega-\nu_\mu)t'$ and neglecting rapidly varying exponentials like $\exp i(\omega+\nu_\mu)t'$. Then $\rho_{ab}{}^{(1)}(a,\mathbf{r},\mathbf{v},t)=$

$$-\tfrac{1}{2}i\left(\frac{\wp}{\hbar}\right)\left[\frac{\lambda_a(\mathbf{r},v,t)}{\gamma_a}\right]\sum_{\mu=1}^{M}\int_{-\infty}^{t}dt'E_\mu(t')\times U_\mu(\mathbf{r}-\mathbf{v}(t-t'))\exp-i(\nu_\mu t+\varphi_\mu(t'))\times\exp[-\gamma_{ab}+i(\nu_\mu-\omega)](t-t'). \quad (37)$$

We also assume that the amplitudes $E_\mu(t')$ and phases $\varphi_\mu(t')$ do not vary much in a time $1/\gamma_{ab}$, so that they can be evaluated at time t. With a change of variable of integration from t' to $\tau'=t-t'$, we find

$$\rho_{ab}{}^{(1)}(a,\mathbf{r},\mathbf{v},t)=-\tfrac{1}{2}i\left(\frac{\wp}{\hbar}\right)\left[\frac{\lambda_a(\mathbf{r},\mathbf{v},t)}{\gamma_a}\right]\sum_{\mu=1}^{M}E_\mu(t)\exp-i(\nu_\mu t+\varphi_\mu(t))\times\int_0^\infty d\tau' U_\mu(\mathbf{r}-\mathbf{v}\tau')\exp-(\gamma_{ab}+i(\omega-\nu_\mu))\tau'. \quad (38)$$

The corresponding contribution to the polarization of the medium is

$$P^{(1)}(a,\mathbf{r},\mathbf{v},t)=-\tfrac{1}{2}i(\wp^2/\hbar)[\lambda_a(\mathbf{r},\mathbf{v},t)/\gamma_a]\times\sum_\mu E_\mu(t)\left\{\exp-i(\nu_\mu t+\varphi_\mu)\int_0^\infty d\tau' U_\mu(\mathbf{r}-\mathbf{v}\tau')\times\exp-(\gamma_{ab}+i(\omega-\nu_\mu))\tau'\right\}+\text{conj.} \quad (39)$$

Let us first assume that the excitation rate density has the form[10]

$$\lambda_\alpha(\mathbf{r},\mathbf{v},t)=W(\mathbf{v})\Lambda_\alpha(\mathbf{r},t)\quad\alpha=a,b, \quad (40)$$

where $W(\mathbf{v})$ is the normalized velocity distribution function and $\Lambda_\alpha(\mathbf{r},t)$ is the number of atoms excited to state α per unit volume and time. Because we are assuming a spatial dependence of the electric field only on z, we may change over from a three- to a one-dimensional description. Then the velocity distribution $W(v)$ refers to the z component v of $\mathbf{v}$, and $\mathbf{r}$ is replaced by z.

It will be noted that the quantity $P(a,z,v,t)$ is proportional to $\Lambda_a(z,t)/\gamma_a$. When we now consider the contribution of atoms excited to the lower maser state b there is a complication which we did not meet in the case of a excitation. Spontaneous decay of atoms in the upper maser level a may be one of the excitation mechanisms for state b. This could be plausibly represented by replacing the excitation rate density $\lambda_b(z,v,t)$ by $\lambda_b(z,v,t)+f\gamma_a\rho_{aa}(z,v,t)$, where f is the branching ratio (decay from a to b)/(total decay from a) and λ_b now describes only "external" excitation processes.

In order to reduce somewhat the complexity of the subsequent equations we will now proceed as if f were zero. The effects of cascade excitation $a\rightarrow b$ will be discussed by an approximate method in Sec. 20. With this simplification it turns out, as one would expect, that for b excitation $P(b,z,v,t)$ is exactly like (39) except for an over-all sign change and interchange of a and b. Hence, the total polarization $P(z,v,t)=P(a,z,v,t)+P(b,z,v,t)$ is proportional to a quantity

$$N(z,t)=[(\Lambda_a(z,t)/\gamma_a)-(\Lambda_b(z,t)/\gamma_b)], \quad (41)$$

which we will call the "excitation density." This is simply the excess density of active atoms in a steady state in the absence of optical oscillations.

The first-order polarization

$$P^{(1)}(z,t)=\wp\int_{-\infty}^{\infty}dvW(v)[\rho_{ab}{}^{(1)}(a,z,v,t)+\rho_{ab}{}^{(1)}(b,z,v,t)+\text{conj.}] \quad (42)$$

is also proportional to $N(z,t)$. For use in Eqs. (13), (14) a spatial Fourier projection on the nth cavity mode is next to be made

$$P_n{}^{(1)}(t)=(2/L)\int_0^L dzP^{(1)}(z,t)U_n(z). \quad (43)$$

The product $U_n(z)U_\mu(z-v\tau')$ which occurs in (43) may be written as

$$\sin K_n z\sin K_\mu(z-v\tau')=\tfrac{1}{2}\cos\{(K_n-K_\mu)z+K_\mu v\tau'\}-\tfrac{1}{2}\cos\{(K_n+K_\mu)z-K_\mu v\tau'\}. \quad (44)$$

The last term will not contribute appreciably to the z integration (43) because the excitation density $N(z,t)$ changes little in an o.f. wavelength. Since the velocity distribution is normally an even function of v only that part of the remainder of (44) which is even in v will contribute to the polarization, i.e.,

$$\tfrac{1}{2}[\cos(K_n-K_\mu)z]\cos Kv\tau',$$

where the subscript μ has been dropped in the last factor since all of the modes considered have very nearly the same wave number $K=\nu/c$.

We find

$$P_n{}^{(1)}(t)=-\tfrac{1}{2}i(\wp^2/\hbar)\sum_\mu E_\mu\exp-i(\nu_\mu t+\varphi_\mu).N_{n-\mu}\times\int_{-\infty}^{\infty}dvW(v)[\mathfrak{D}(\omega-\nu_\mu+Kv)]+\text{conj.}, \quad (45)$$

where

$$\mathfrak{D}(\omega)=1/(\gamma_{ab}+i\omega) \quad (46)$$

[10] It would be easy to modify the theory to allow the atoms excited to state b to have a different velocity distribution from those excited to state a.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is a convenient abbreviation for a frequently occurring denominator and where

$$N_{n-\mu}(t)=\frac{1}{L}\int_0^L dz N(z,t)\cos\left[(n-\mu)\frac{\pi z}{L}\right] \quad (47)$$

is a spatial Fourier component of the excitation density $N(z,t)$. It should be noted that (45) has a very simple interpretation in terms of Doppler shifts of the atomic transition frequencies by Kv due to the atomic motion. This simplicity will be lost when nonlinear effects are considered.

For the following detailed calculations a Maxwellian distribution

$$W(v)=(u\pi^{1/2})^{-1}\exp-(v^2/u^2) \quad (48)$$

will be assumed. The speed parameter u is related to an effective temperature T by the equation

$$\tfrac{1}{2}mu^2=k_BT, \quad (49)$$

where m is the atomic mass and k_B is the Boltzmann constant. If it should develop that a Maxwellian distribution is not realized in practice, some obvious changes in the later work can be made.

With Eq. (48) the integration over v may profitably be done on (39) before that over τ', and we find

$$P_n^{(1)}(t)=-\tfrac{1}{2}(\wp^2/\hbar Ku)\times\Big[\sum_{\mu=1}^{M}E_\mu(t)\exp-i(\nu_\mu t+\varphi_\mu(t))\times N_{n-\mu}(t)Z(\nu_\mu-\omega)+\text{conj.}\Big], \quad (50)$$

where $Z(\nu-\omega)$ is an abbreviation for

$$Z(\nu-\omega,\gamma_{ab},Ku)=iKu\int_0^\infty d\tau\exp\left[i(\nu-\omega)\tau-\gamma_{ab}\tau-\tfrac{1}{4}K^2u^2\tau^2\right], \quad (51)$$

which is a complex function well known in the theory of Doppler broadening.[11] The function Z is, in fact, a function of a single complex variable ζ

$$Z(\zeta)=2i\int_{-\infty}^{i\zeta}dt\exp-(t^2+\zeta^2), \quad (52)$$

where

$$\zeta=\xi+i\eta, \quad (53)$$

with

$$\xi=(\nu-\omega)/Ku, \quad (54)$$

and

$$\eta=\gamma_{ab}/Ku. \quad (55)$$

It is fortunate that extensive tables[12] of the real part Z_r and the imaginary part Z_i of $Z(\xi+i\eta)$ are now available.

$P_n^{(1)}(t)$ is a linear function of the complex electric fields $E_\mu(t)\exp-i\nu_\mu t-i\varphi_\mu(t)$ of the cavity modes and, apart from amplitude modulation arising from a possible slow time variation of the excitation density $N(z,t)$ contains the same frequencies as the cavity field.

To determine amplitude and frequency (or phase) of the oscillations, we write out the contributions of $P_n^{(1)}$ to C_n and S_n of Eqs. (13), (14). These are

$$S_n^{(1)}=-(\wp^2/\hbar Ku)\bar{N}Z_i(\nu_n-\omega)E_n, \quad (56)$$

$$C_n^{(1)}=-(\wp^2/\hbar Ku)\bar{N}Z_r(\nu_n-\omega)E_n, \quad (57)$$

where

$$\bar{N}=N_0(t)=(1/L)\int_0^L dzN(zt), \quad (58)$$

which will be called the "excitation," is the average of the excitation density over the cavity. We have now reverted to the notation of Eq. (51) for the Z function, but to shorten equations have dropped the parameters γ_{ab} and Ku which appear as arguments in (51).

In this approximation, without nonlinear terms, we can only hope to obtain the condition for starting of oscillations and their frequency at threshold. Furthermore, if the conditions are such that several modes can oscillate, they do so independently of each other and hence can be considered separately. The amplitude equation (14) gives

$$\dot{E}_n=-\tfrac{1}{2}(\nu/Q_n)E_\mu-\tfrac{1}{2}(\nu/\epsilon_0)S_n^{(1)} \quad (59)$$

or for a steady state, for which $\dot{E}_n=0$,

$$(\wp^2/\epsilon_0\hbar Ku)\bar{N}Z_i(\nu_n-\omega)=1/Q_n. \quad (60)$$

To first order in $\eta=\gamma_{ab}/Ku$, we have

$$Z(\xi,\eta)\simeq(1-2i\eta\xi)\left[-2\int_0^\xi e^{x^2}dx+i\pi^{1/2}\right]e^{-\xi^2}-2i\eta. \quad (61)$$

Hence for pure Doppler line shape the condition (60) for the onset of oscillations in the nth mode may be written as

$$2\pi^{1/2}[e^2/(4\pi\epsilon_0\hbar u)]\times(\wp/e)^2\lambda\bar{N}\exp-(\nu_n-\omega)^2/(Ku)^2=1/Q_n, \quad (62)$$

where $\lambda=2\pi/K$ is the wavelength. In these units,

$$e^2/(4\pi\epsilon_0\hbar c)\approx 1/137 \quad (63)$$

is the fine structure constant, so the left-hand side of (62) is the product of this and four other dimensionless factors: c/u, $2\pi^{1/2}$, $\exp-(\nu_n-\omega)^2/(Ku)^2$ and $(\wp/e)^2\lambda\bar{N}$ which is the net number of active atoms in a cylinder of cross sectional area $(\wp/e)^2$ and length λ. As the frequency detuning $\nu_n-\omega$ increases, the excitation $\bar{N}$ required to initiate oscillations increases in proportion to $\exp+(\nu_n-\omega)^2/(Ku)^2$. The frequency of oscillation is

[11] M. Born, *Optik* (Julius Springer-Verlag, Berlin, 1933), pp. 482–486.

[12] B. D. Fried and S. D. Conte, *The Plasma Dispersion Function (Hilbert Transform of the Gaussian)* (Academic Press, Inc., New York, 1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

determined by Eq. (13) in which we may set $\dot{\varphi}_n=0$ without loss of generality. Using Eq. (57), we find

$$\nu_n=\Omega_n+(\nu/2)(\wp^2\bar{N}/\epsilon_0\hbar Ku)Z_r(\nu_n-\omega). \quad (64)$$

It is convenient to use the threshold condition (60) to express $\bar{N}$ in terms of Q_n. We find

$$\nu_n-\Omega_n=\tfrac{1}{2}(\nu/Q_n)[Z_r(\nu_n-\omega)/Z_i(\nu_n-\omega)] \quad (65)$$

or in the approximation $\gamma_{ab}/Ku\ll 1$

$$\nu_n-\Omega_n=-\left(\frac{1}{\pi^{1/2}}\right)\left(\frac{\nu}{Q_n}\right)\int_0^{\xi_n}dx e^{x^2}. \quad (66)$$

If the integral (66) is expanded to first order in $\xi_n=(\nu_n-\omega)/Ku$ we obtain

$$(\nu_n-\Omega_n)/(\omega-\nu_n)\approx(1/\pi^{1/2})(\nu/QKu)=\mathcal{S}, \quad (67)$$

which implies "linear pulling," i.e., the detuning of ν_n from the cavity frequency Ω_n is proportional to the amount $\omega-\nu_n$ by which the oscillator frequency ν_n is removed from the atomic resonance frequency ω. The right-hand side of (67) ($\mathcal{S}$="stabilization factor"[13]) is about 1/800 for typical values of the parameters used in Sec. 2.

The more accurate expression (65) indicates a "nonlinear pulling" such that the oscillator frequency is nearer to the atomic frequency than it would be for linear pulling. To the next order in $(\Omega_n-\omega)/Ku$, but still for $\gamma_{ab}\ll Ku$, we find

$$(\nu_n-\Omega_n)/(\omega-\nu_n)\simeq[1+\tfrac{1}{3}(\Omega_n-\omega)^2/(Ku)^2]\mathcal{S}. \quad (68)$$

7. THIRD-ORDER TERMS

We now carry out a similar calculation for the third-order quantity $P_n^{(3)}(t)$ using Eqs. (8), (23), (30), (31) and (32). The integration of $\lambda_n(z,v,t_0)\rho_{ab}^{(3)}(a,z,t_0,v,t)$ over times t_0 of excitation involves integrals like

$$\int_{-\infty}^{t}dt_0\int_{t_0}^{t}dt'\int_{t_0}^{t'}dt''\int_{t_0}^{t''}dt'''F(t,t',t'',t''')\exp-\gamma_a(t'''-t_0),$$

which by a repeated interchange of orders of integration can be reduced to

$$(1/\gamma_a)\int_{-\infty}^{t}dt'\int_{-\infty}^{t'}dt''\int_{-\infty}^{t''}dt'''F(t,t',t'',t''').$$

Again, we keep only exponential factors in the time integration which are able to have resonance, and find after changes of variables

$$\tau'=t-t', \quad \tau''=t'-t'', \quad \tau'''-t''-t''' \quad (69)$$

and some algebraic manipulations

$$\begin{aligned}\rho_{ab}^{(3)}(z,v,t)=\tfrac{1}{8}i\wp^3\hbar^{-3}N(z,t)\sum_\mu\sum_\rho\sum_\sigma E_\mu E_\rho E_\sigma\Big\{&[\exp-i(\nu_\mu t+\varphi_\mu)+i(\nu_\rho t+\varphi_\rho)-i(\nu_\sigma t+\varphi_\sigma)]\\
&\times\int_0^\infty d\tau'\int_0^\infty d\tau''\int_0^\infty d\tau'''U_\mu(z-v\tau')U_\rho(z-v\tau'-v\tau'')U_\sigma(z-v\tau'-v\tau''-v\tau''')\\
&\times\{\exp-[(\gamma_{ab}-i\nu_\mu+i\nu_\rho-i\nu_\sigma+i\omega)\tau'+(\gamma_a+i\nu_\rho-i\nu_\sigma)\tau''+(\gamma_{ab}+i\omega-i\nu_\sigma)\tau'''\}\\
&+[\exp-i(\nu_\mu t+\varphi_\mu)-i(\nu_\rho t+\varphi_\rho)+i(\nu_\sigma t+\varphi_\sigma)]\\
&\times\int_0^\infty d\tau'\int_0^\infty d\tau''\int_0^\infty d\tau'''U_\mu(z-v\tau')U_\rho(z-v\tau'-v\tau'')U_\sigma(z-v\tau'-v\tau''-v\tau''')\\
&\times[\exp-[(\gamma_{ab}-i\nu_\mu-i\nu_\rho+i\nu_\sigma+i\omega)\tau'+(\gamma_a-i\nu_\rho+i\nu_\sigma)\tau''+(\gamma_{ab}-i\omega+i\nu_\sigma)\tau'''\Big\}\end{aligned}$$

$$+\text{same with } a \text{ and } b \text{ interchanged.} \quad (70)$$

In calculation of the Fourier projection $P_n^{(3)}(t)$ integrals of the form

$$\left(\frac{2}{L}\right)\int_0^L dz N(zt)U_n(z)U_\mu(z-v\tau')U_\rho(z-v\tau'-v\tau'')U_\sigma(z-v\tau'-v\tau''-v\tau''')$$

appear. The product of the four sine functions can be reduced to

$$\tfrac{1}{8}[\cos(K_\rho-K_\sigma+K_\mu-K_n)z\cos Kv(\tau'''-\tau')+\cos(K_\rho-K_\sigma-K_\mu+K_n)z\cos Kv(\tau'''+\tau')$$
$$+\cos(K_\rho+K_\sigma-K_\mu-K_n)z\cos Kv(\tau'+2\tau''+\tau''')].$$

[13] The term "stabilization factor" (≈cavity band width/atom bandwidth) was previously used in a theory of the ammonia beam maser (Ref. 8, p. 460) where its numerical value was large compared to unity.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

apart from rapidly oscillating terms and those odd in v. The integration over v may now be carried out, and we find

$$P_n^{(3)}(t)=\tfrac{1}{32}i\wp^4\hbar^{-3}\sum_\mu\sum_\rho\sum_\sigma E_\mu E_\rho E_\sigma\Big\{[\exp-i(\nu_\mu t+\varphi_\mu)+i(\nu_\rho t+\varphi_\rho)-i(\nu_\sigma t+\varphi_\sigma)]$$
$$\times\int_0^\infty d\tau'\int_0^\infty d\tau''\int_0^\infty d\tau'''[N_{(\rho-\sigma+\mu-n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'''-\tau')^2+N_{(\rho-\sigma-\mu+n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'''+\tau')^2$$
$$+N_{(\rho+\sigma-\mu-n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'+2\tau''+\tau''')^2][\exp-(\gamma_{ab}-i\nu_\mu+i\nu_\rho-i\nu_\sigma+i\omega)\tau'$$
$$-(\gamma_a+i\nu_\rho-i\nu_\sigma)\tau''-(\gamma_{ab}+i\omega-i\nu_\sigma)\tau'''][\exp-i(\nu_\mu t+\varphi_\mu)-i(\nu_\rho t+\varphi_\rho)+i(v_\sigma t+\varphi_\sigma)]$$
$$\times\int_0^\infty d\tau'\int_0^\infty d\tau''\int_0^\infty d\tau'''[N_{(\rho-\sigma+\mu-n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'''-\tau')^2+N_{(\rho-\sigma-\mu+n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'''+\tau')^2$$
$$+N_{(\rho+\sigma-\mu-n)}\exp-\tfrac{1}{4}(Ku)^2(\tau'+2\tau''+\tau''')^2][\exp-(\gamma_{ab}-i\nu_\mu-i\nu_\rho+i\nu_\sigma+i\omega)\tau'-(\gamma_a-i\nu_\rho+i\nu_\sigma)\tau''$$
$$-(\gamma_{ab}-i\omega+i\nu_\sigma)\tau''']\Big\}+\text{same with }\gamma_a\text{ and }\gamma_b\text{ interchanged}+\text{complex conjugate.}\quad(71)$$

One of the most important characteristics of the third-order polarization is that it has constituents which oscillate at all possible frequencies $\nu_\mu\mp\nu_\rho\pm\nu_\sigma$. The combination of signs appearing here is correlated with use of the rotating wave approximation. Terms with frequencies such as $\nu_\mu+\nu_\rho+\nu_\sigma$ (near the third harmonic) are thereby neglected.

The formidable expression (71) can be simplified in either of two limiting cases: (a) no atomic motion ($u=0$), or (b) "Doppler limit," i.e., Ku much larger than γ_{ab} and various frequency differences such as $\nu_\mu+\nu_\rho-2\nu_\sigma$, etc. In the absence of atomic motion, one finds after some rewriting

$$P_n^{(3)}(t)=\tfrac{1}{32}i\hbar^{-3}\wp^4\sum_\mu\sum_\rho\sum_\sigma E_\mu E_\rho E_\sigma[\exp-i(\nu_\mu t+\varphi_\mu)+i(\nu_\rho t+\varphi_\rho)-i(\nu_\sigma t+\varphi_\sigma)]$$
$$\times[N_{(\rho-\sigma+\mu-n)}+N_{(\rho-\sigma-\mu+n)}+N_{(\rho+\sigma-\mu-n)}]\cdot\mathfrak{D}(\nu_\rho-\nu_\mu-\nu_\sigma+\omega)$$
$$\times[\mathfrak{D}_a(\nu_\rho-\nu_\sigma)+\mathfrak{D}_b(\nu_\rho-\nu_\sigma)][\mathfrak{D}(\omega-\nu_\sigma)+\mathfrak{D}(\nu_\rho-\omega)]+\text{conjugate},\quad(72)$$

where $\mathfrak{D}(\omega)$ was defined by Eq. (46) and

$$\mathfrak{D}_\alpha(\omega)=1/(\gamma_\alpha+i\omega),\quad\alpha=a,b.\quad(73)$$

The "Doppler limit" is appropriate for many possible gaseous optical masers, and for most of the remainder of this paper we will be dealing only with this case. Then

$$\exp-\tfrac{1}{4}(Ku)^2(\tau'''-\tau')^2$$

acts like a delta function of $\tau'''-\tau'$ and the integration over τ''' can be done in the form

$$\int_0^\infty d\tau'\int_0^\infty d\tau'''G(\tau',\tau''')\exp-\tfrac{1}{4}(Ku)^2(\tau'''-\tau')^2\simeq\frac{2\pi^{1/2}}{Ku}\int_0^\infty d\tau' G(\tau',\tau').\quad(74)$$

The other Gaussian factors do not have their full peaks in the range of integration, and give contributions which we neglect because they lead to expressions with higher powers of Ku in the denominator. Then, after performing the simple integrations over τ' and τ'', we find

$$P_n^{(3)}(t)=\tfrac{1}{32}i\pi^{1/2}[\wp^4/(\hbar^3Ku)]\sum_\mu\sum_\rho\sum_\sigma E_\mu E_\rho E_\sigma N_{(\rho-\sigma+\mu-n)}\{\mathfrak{D}(\omega-\tfrac{1}{2}\nu_\mu+\tfrac{1}{2}\nu_\rho-\nu_\sigma)[\mathfrak{D}_a(\nu_\rho-\nu_\sigma)+\mathfrak{D}_b(\nu_\rho-\nu_\sigma)]$$
$$\times[\exp-i(\nu_\mu-\nu_\rho+\nu_\sigma)t-i(\varphi_\mu-\varphi_\rho+\varphi_\sigma)]+\mathfrak{D}(-\tfrac{1}{2}\nu_\mu-\tfrac{1}{2}\nu_\rho+\nu_\sigma)[\mathfrak{D}_a(\nu_\sigma-\nu_\rho)+\mathfrak{D}_b(\nu_\sigma-\nu_\rho)]$$
$$\times[\exp[-i(\nu_\mu+\nu_\rho-\nu_\sigma)t-i(\varphi_\mu+\varphi_\rho-\varphi_\sigma)]]\}+\text{complex conjugate.}\quad(75)$$

By interchanging ρ and σ in the second group of terms, this may be written more compactly as

$$P_n^{(3)}(t)=\tfrac{1}{32}i\pi^{1/2}[\wp^4/(\hbar^3Ku)]\sum_\mu\sum_\rho\sum_\sigma E_\mu E_\rho E_\sigma[\exp[-i(\nu_\mu-\nu_\rho+\nu_\sigma)t-i(\varphi_\mu-\varphi_\rho+\varphi_\sigma)]]$$
$$\times[N_{(\rho-\sigma+\mu-n)}\mathfrak{D}(\omega-\tfrac{1}{2}\nu_\mu+\tfrac{1}{2}\nu_\rho-\nu_\sigma)+N_{(\sigma-\rho+\mu-n)}\mathfrak{D}(-\tfrac{1}{2}\nu_\mu-\tfrac{1}{2}\nu_\sigma+\nu_\rho)][\mathfrak{D}_a(\nu_\rho-\nu_\sigma)+\mathfrak{D}_b(\nu_\rho-\nu_\sigma)]$$
$$+\text{complex conjugate.}\quad(76)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

8. SINGLE FREQUENCY OPERATION

For this case, the triple summation over μ, ρ, and σ reduces to a single term

$$P_n^{(3)}(t)=\tfrac{1}{16}i\pi^{1/2}\bar{N}[\wp^4/(\hbar^3\gamma_a\gamma_b Ku)]\times\gamma_{ab}[\mathfrak{D}(\omega-\nu_n)+\mathfrak{D}(0)]E_n^{(3)}\exp-i(\nu_n t+\varphi_n)+\text{complex conjugate.} \quad (77)$$

For use in Eqs. (13), (14) we need the in-phase and quadrature components of this

$$C_n^{(3)}(t)=\tfrac{1}{8}\pi^{1/2}\bar{N}[\wp^4/(\hbar^3\gamma_a\gamma_b Ku)]\times\gamma_{ab}(\omega-\nu_n)\mathfrak{L}(\omega-\nu_n)E_n^3 \quad (78)$$

and

$$S_n^{(3)}(t)=\tfrac{1}{8}\pi^{1/2}\bar{N}[\wp^4/(\hbar^3\gamma_a\gamma_b Ku)]\times[1+\gamma_{ab}^2\mathfrak{L}(\omega-\nu_n)]E_n^3, \quad (79)$$

where the Lorentzian function is denoted by

$$\mathfrak{L}(\omega-\nu)=[\gamma_{ab}^2+(\omega-\nu)^2]^{-1}. \quad (80)$$

The amplitude determining Eq. (13) is then of the form[14]

$$\dot{E}_n=\alpha_n E_n-\beta_n E_n^3, \quad (81)$$

where

$$\alpha_n=-\tfrac{1}{2}(\nu/Q_n)+\tfrac{1}{2}\nu\bar{N}[\wp^2/(\epsilon_0\hbar Ku)]Z_i(\nu_n-\omega), \quad (82)$$

$$\beta_n=(\nu/16)\pi^{1/2}\bar{N}[\wp^4/(\epsilon_0\hbar^3\gamma_a\gamma_b Ku)]\times[1+\gamma_{ab}^2\mathfrak{L}(\nu_n-\omega)]. \quad (83)$$

It is useful to employ the starting condition (60) to express the coefficients α_n, β_n (and others which appear later) in terms of a ratio

$$\mathfrak{N}=\bar{N}/\bar{N}_T \quad (84)$$

called the "relative excitation" where $\bar{N}_T$ is the excitation required for threshold oscillations when the cavity frequency Ω_n is tuned to the peak ω of the atomic resonance curve. We find

$$\alpha_n=\tfrac{1}{2}(\nu/Q_n)\{[Z_i(\nu_n-\omega)/Z_i(0)]\mathfrak{N}-1\} \quad (85)$$

and

$$\beta_n=\tfrac{1}{16}\pi^{1/2}(\nu/Q_n)[\mathfrak{N}\wp^2/(\hbar^2\gamma_a\gamma_b Z_i(0))]\times[1+\gamma_{ab}^2\mathfrak{L}(\nu_n-\omega)]. \quad (86)$$

A stable steady state occurs for an intensity of oscillations

$$E_n^2=\alpha_n/\beta_n, \quad (87)$$

which is easily related to the relative excitation with the help of Eqs. (85) and (86).

The frequency determining equation (again $\dot{\varphi}_n$ may be taken zero) is

$$\nu_n=\Omega_n-\tfrac{1}{2}(\nu/(\epsilon_0 E_n))[C_n^{(1)}+C_n^{(3)}]. \quad (88)$$

Since the frequency of oscillation ν_n will differ little from the cavity resonance frequency Ω_n, the right side of Eq. (88) may, to a sufficiently good approximation, be evaluated for $\nu_n=\Omega_n$. We may then write

$$\nu_n=\Omega_n+\sigma_n+\rho_n E_n^2, \quad (89)$$

where

$$\sigma_n=\tfrac{1}{2}\nu[\wp^2\bar{N}/(\epsilon_0\hbar Ku)]Z_r(\Omega_n-\omega) \quad (90)$$

and

$$\rho_n=+\tfrac{1}{16}\pi^{1/2}\nu[\wp^4\bar{N}/(\epsilon_0\hbar^2\gamma_a\gamma_b Ku)]\times\gamma_{ab}(\Omega_n-\omega)\mathfrak{L}(\Omega_n-\omega) \quad (91)$$

or expressing these coefficients in terms of the relative excitation (84)

$$\sigma_n=\tfrac{1}{2}(\nu/Q_n)\mathfrak{N}Z_r(\Omega_n-\omega)/Z_i(0) \quad (92)$$

and

$$\rho_n=+\tfrac{1}{16}\pi^{1/2}(\nu/Q_n)\mathfrak{N}[\wp^2/(\hbar^2\gamma_a\gamma_b Z_i(0))]\times\gamma_{ab}(\Omega_n-\omega)\mathfrak{L}(\Omega_n-\omega) \quad (93)$$

so that

$$\rho_n E_n^2=\rho_n\alpha_n/\beta_n=\tfrac{1}{2}(\nu/Q_n)\{[Z_i(\Omega_n-\omega)/Z_i(0)]\mathfrak{N}-1\}\gamma_{ab}(\Omega_n-\omega)/[2\gamma_{ab}^2+(\Omega_n-\omega)^2]. \quad (94)$$

The frequency now depends on the relative excitation (and hence on the power level) as well as on the detuning, i.e., there is frequency "pushing" as well as "pulling."

Equations (89), (92), and (94) indicate that for

$$(\Omega_n-\omega)^2+2\gamma_{ab}^2<-\gamma_{ab}(\Omega_n-\omega)Z_i(\Omega_n-\omega)/Z_r(\Omega_n-\omega) \quad (95)$$

an increase of excitation should move the frequency in the direction from ω toward Ω_n. For small detuning the right side of (95) is approximately $\frac{1}{2}\pi^{1/2}\gamma_{ab}Ku$.

The dependence of power level on excitation and detuning is given by Eq. (87). Using the value of $Z_i(\xi_n,\eta)$ for $\mathfrak{N}=0$, this equation may be written approximately as

$$(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)=8\left\{\left[\exp-\frac{(\Omega_n-\omega)^2}{(Ku)^2}\right]-\mathfrak{N}^{-1}\right\}\Big/[1+\gamma_{ab}^2\mathfrak{L}(\Omega_n-\omega)]. \quad (96)$$

This expression agrees with the linear approximation (62) in predicting threshold for relative excitation

$$\mathfrak{N}=\bar{N}/\bar{N}_T=\exp(\Omega_n-\omega)^2/(Ku)^2 \quad (97)$$

when there is detuning. Because of its derivation from a third order perturbation theory, Eq. (96) should not be trusted unless it predicts a value of the "saturation

[14] It is easy to see that the coefficient α_n is simply related to the gain (negative absorption) coefficient of the medium at frequency ν_n for small signals. However, the gain coefficient for a strong signal cannot be safely inferred from Eq. (81), since standing rather than traveling waves were assumed in its derivation. A theory of a traveling wave maser along the lines of this paper will be given later.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

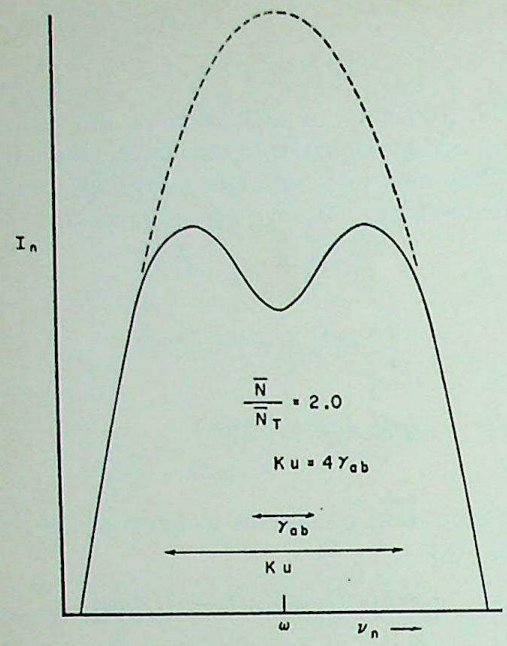


FIG. 2. Relative intensity of oscillation as a function of detuning. The solid curve, drawn for parameters $\bar{N}=2\bar{N}_T$ and $Ku=4\gamma_{ab}$ represents Eq. (96). The dotted curve indicates the Doppler gain profile of the numerator of (96).

parameter"[15]

$$I_n=\tfrac{1}{2}(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b) \tag{98}$$

much less than unity.

The numerator in Eq. (96) has a peak for resonant tuning $\Omega_n=\omega$, but the denominator, which comes from the nonlinear term involving the coefficient β_n also has a peak when $\Omega_n=\omega$. Under certain conditions, the over-all curve of $E_n^2=\alpha_n/\beta_n$ versus detuning $\Omega_n-\omega$ should have a flattened peak at resonance, or even a dip between two maxima. The condition for the appearance of two maxima is

$$(\gamma_{ab}/Ku)^2<\tfrac{1}{2}\{1-\exp-[(\Omega_n{}^*-\omega)^2/(Ku)^2]\}, \tag{99}$$

where $\Omega_n{}^*-\omega$ is the detuning required to stop oscillations at the given level of excitation. The double peak[16] (see Fig. 2) should thus be seen somewhat above threshold, i.e., for relative excitation

$$\mathfrak{N}>1/[1-2(\gamma_{ab}/(Ku))^2]. \tag{100}$$

Under this degree of relative excitation, the electric field at central tuning is given by

$$I_n=\tfrac{1}{2}(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)\approx 4(\gamma_{ab}/(Ku))^2\ll 1 \tag{101}$$

so that the neglect of higher orders of perturbation theory should not be too serious, provided, as we assume, that $\gamma_{ab}\ll Ku$.

[15] The significance of this quantity is shown more clearly in Sec. 18.

[16] This dip has recently been observed. R. A. McFarlane, W. R. Bennett, and W. E. Lamb, Appl. Phys. Letters **2**, 189 (1963); A. Szöke and A. Javan, Phys. Rev. Letters **10**, 521 (1963).

TABLE I. This shows the twenty-seven possible values of the summation indices μ, ρ, σ which appear in Eq. (76) for "three"-mode oscillation. The fourth column gives the corresponding frequencies $\nu_\mu-\nu_\rho+\nu_\sigma$. The last column contains numerical values for the amount by which these frequencies exceed ν_1. We have taken $\nu_2-\nu_1=150$ Mc/sec and $\nu_3-\nu_2=151$ Mc/sec in order to simulate (but greatly to exaggerate) nonlinear pulling effects.

μ	ρ	σ	$\nu_\mu-\nu_\rho+\nu_\sigma$	Typical $\delta\nu$ (Mc/sec)
1	1	1	ν_1	0
1	1	2	ν_2	150
1	1	3	ν_3	301
1	2	1	$2\nu_1-\nu_2$	−150
1	2	2	ν_1	0
1	2	3	$\nu_1+\nu_3-\nu_2$	151
1	3	1	$2\nu_1-\nu_3$	−301
1	3	2	$\nu_1+\nu_2-\nu_3$	−151
1	3	3	ν_1	0
2	1	1	ν_2	150
2	1	2	$2\nu_2-\nu_1$	300
2	1	3	$\nu_2+\nu_3-\nu_1$	451
2	2	1	ν_1	0
2	2	2	ν_2	150
2	2	3	ν_3	301
2	3	1	$\nu_1+\nu_2-\nu_3$	−151
2	3	2	$2\nu_2-\nu_3$	−1
2	3	3	ν_2	150
3	1	1	ν_3	301
3	1	2	$\nu_2+\nu_3-\nu_1$	451
3	1	3	$2\nu_3-\nu_1$	602
3	2	1	$\nu_1+\nu_3-\nu_2$	151
3	2	2	ν_3	301
3	2	3	$2\nu_3-\nu_2$	452
3	3	1	ν_1	0
3	3	2	ν_2	150
3	3	3	ν_3	301

The physical interpretation of the dip is discussed in Secs. 17 and 18 from several points of view.

9. MULTIPLE MODE OPERATION

As the excitation is increased beyond that required for threshold of single frequency oscillation, other frequencies appear in the output of an optical maser oscillator. We wish to use the expression (76) for the nonlinear polarization in the electromagnetic field equations (13), (14) in order to account for the observed phenomena.

The first theory of an oscillator capable of multifrequency operation was given by van der Pol[17] in 1921–22. The necessary nonlinear features were provided by cubic terms in the current-voltage characteristic of a triode vacuum tube. The tank had two *R-L-C* circuits with resonance frequencies Ω_1 and Ω_2. Van der Pol found that steady oscillations could occur only at frequencies near Ω_1 or Ω_2, but that simultaneous steady operation at the two frequencies was impossible. There were hysteresis phenomena, i.e., the choice of steady state of oscillation depended on the past history of the circuit parameters.

Multicavity magnetrons provide very important

[17] B. van der Pol, Phil. Mag. **43**, 700 (1922) and a review article Proc. Inst. Radio Engrs. **22**, 1051 (1934).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. This table gives various quantities needed for the evaluation of Eq. (76) for three-frequency oscillation as described in the text.

μ	ρ	σ	n	$\nu_\rho-\nu_\sigma$	$\omega-\frac{1}{2}\nu_\mu+\frac{1}{2}\nu_\rho-\nu_\sigma$	$-\frac{1}{2}\nu_\mu-\frac{1}{2}\nu_\sigma+\nu_\rho$	$\nu_\mu-\nu_\rho+\nu_\sigma$	$\rho-\sigma+\mu-n$	$\sigma-\rho+\mu-n$
1	1	1	1	0	$\omega-\nu_1$	0	ν_1	0	0
1	2	2	1	0	$\omega-\nu_{12}$	$\frac{1}{2}\Delta$	ν_1	0	0
1	3	3	1	0	$\omega-\nu_2$	Δ	ν_1	0	0
2	2	1	1	Δ	$\omega-\nu_1$	$\frac{1}{2}\Delta$	ν_1	2	0
3	3	1	1	2Δ	$\omega-\nu_1$	Δ	ν_1	4	0
2	3	2	1	Δ	$\omega-\nu_{12}$	Δ	$2\nu_2-\nu_3$	2	0
2	2	2	2	0	$\omega-\nu_2$	0	ν_2	0	0
1	1	2	2	$-\Delta$	$\omega-\nu_2$	$-\frac{1}{2}\Delta$	ν_2	-2	0
2	1	1	2	0	$\omega-\nu_{12}$	$-\frac{1}{2}\Delta$	ν_2	0	0
2	3	3	2	0	$\omega-\nu_{23}$	$\frac{1}{2}\Delta$	ν_2	0	0
3	3	2	2	Δ	$\omega-\nu_2$	$\frac{1}{2}\Delta$	ν_2	2	0
1	2	3	2	$-\Delta$	$\omega-\nu_{23}$	"0"	$\nu_1+\nu_3-\nu_2$	-2	0
3	2	1	2	Δ	$\omega-\nu_{12}$	"0"	$\nu_1+\nu_3-\nu_2$	2	0
3	3	3	3	0	$\omega-\nu_3$	0	ν_3	0	0
1	1	3	3	-2Δ	$\omega-\nu_3$	$-\Delta$	ν_3	-4	0
2	2	3	3	$-\Delta$	$\omega-\nu_3$	$-\frac{1}{2}\Delta$	ν_3	-2	0
3	1	1	3	0	$\omega-\nu_2$	$-\Delta$	ν_3	0	0
3	2	2	3	0	$\omega-\nu_{23}$	$-\frac{1}{2}\Delta$	ν_3	0	0
2	1	2	3	$-\Delta$	$\omega-\nu_{23}$	$-\Delta$	$2\nu_2-\nu_1$	-2	0

examples of oscillators capable of multifrequency operation, and here again the normal pattern is that an oscillation existing at one frequency tends to suppress one at another frequency.

There is, of course, a very close connection between van der Pol's work and ours. Where he dealt with a nonlinear triode characteristic for a vacuum tube, we are concerned with the nonlinear response of an assembly of atomic systems which obey the laws of quantum mechanics. A one-to-one correspondence can be set up between the two problems. The Fabry-Perot cavity modes correspond to the resonant constituents of van der Pol's plate circuit. As we will see, however, the effective tube characteristics of the optical medium differ qualitatively from that assumed by van der Pol, and hence the optical maser behaves in a very different fashion with respect to multifrequency operation than the oscillator in van der Pol's original model.

Rather than deal next with the case of two-frequency operation, it will perhaps save space to consider first the more general case of three-frequency operation. Probably most of the interesting phenomena for optical gaseous masers can be understood without dealing explicitly with more than three frequencies. After the more general equations have been obtained, we can easily drop terms and discuss two frequency oscillation as a special and simpler case.

We consider the expression (76) for the third-order polarization, in which the indices μ, ρ, σ, and n can each take on values 1, 2, and 3. It is useful to have the ingredients of the summands in tabular form. The entries in the fourth column of Table I give the frequencies of the various summands in $P_n^{(3)}$ identified by the μ, ρ, and σ values in the first three columns. It will be noted that besides the three frequencies ν_1, ν_2, and ν_3 assumed in the cavity excitation, there are nine additional frequencies present in the polarization of the medium. Hence there must be fields in the cavity at these new frequencies, and the desired self-consistency for three-frequency operation is in jeopardy. However, under certain conditions which will be determined later, the fields at the new frequencies do not produce appreciable effects, even though the frequencies lie close to cavity resonance, so that the calculation can be made as planned.

As is already apparent from the single-frequency case, the oscillation frequencies ν_n are typically very close to the cavity frequencies Ω_n, which are equally spaced and separated by

$$\Delta \approx 150 \text{ Mc/sec.} \tag{102}$$

Hence three of the new frequencies: $2\nu_2-\nu_3$, $\nu_1+\nu_3-\nu_2$, and $2\nu_2-\nu_1$ are very close to the three main frequencies ν_1, ν_2, and ν_3, respectively, and the corresponding terms are carried along in the calculations. The remaining frequencies can be ignored as long as the oscillator is appreciably below threshold for four-frequency operation.

Any pulling of ν_n from Ω_n is typically measured in kc/sec and hence the ν_n are not detuned from Ω_n by an appreciable fraction either of the cavity bandwidths ν/Q_n which are about 1 Mc/sec, or of the radiative decay constants γ_a, γ_b, or γ_{ab} which may be 10 Mc/sec or more. The entries in columns 5, 6, and 7 of Table II occur in the frequency denominators in Eq. (76) and are there to be combined with imaginary numbers $-i\gamma_a$, $-i\gamma_b$, or $-i\gamma_{ab}$. We have neglected the small terms arising from frequency pulling. Thus $\nu_2-\nu_1$ is freely replaced by Δ of Eq. (102), etc. A symbol like ν_{12} denotes a frequency halfway between ν_1 and ν_2, etc.,

$$\nu_{12}=\tfrac{1}{2}(\nu_1+\nu_2), \text{ etc.} \tag{103}$$

The entries in the last two columns of Table II are the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

integers characterizing the spatial Fourier components (47) of the excitation density which are needed for the evaluation of the contribution to (76) arising from the indices μ, ρ, σ, and n.

The combinations (μ,ρ,σ) which contribute third order polarizations at frequency ν_1 are (1,1,1), (1,2,2), (1,3,3), (2,2,1), and (3,3,1) while (2,3,2) gives a contribution near ν_1. Each summand has a product of two-resonance denominators. The dominant terms are (1,1,1), (1,2,2), and (1,3,3) since they do not necessarily contain inverse powers of Δ. The other terms have at least one power of Δ in the denominator. All terms will ultimately be expanded to order $1/\Delta^2$.

We find

$$P_1^{(3)}(t)=\tfrac{1}{16}i\pi^{1/2}[\wp^4/\hbar^3Ku)][\{E_1^3\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b))(\mathfrak{D}(\omega-\nu_1)+\mathfrak{D}(0))$$
$$+E_1E_2^2\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b))(\mathfrak{D}(\omega-\nu_{12})+\mathfrak{D}(\Delta/2))+E_1E_3^2\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b))(\mathfrak{D}(\omega-\nu_2)+\mathfrak{D}(\Delta))$$
$$+E_2^2E_1[\tfrac{1}{2}N_2\mathfrak{D}(\omega-\nu_1)+\tfrac{1}{2}\bar{N}\mathfrak{D}(\Delta/2)][\mathfrak{D}_a(\Delta)+\mathfrak{D}_b(\Delta)]+E_3^2E_1[\tfrac{1}{2}N_4\mathfrak{D}(\omega-\nu_1)+\tfrac{1}{2}\bar{N}\mathfrak{D}(\Delta)]$$
$$\times[\mathfrak{D}_a(2\Delta)+\mathfrak{D}_b(2\Delta)]\}\exp-i(\nu_1t+\varphi_1)+\tfrac{1}{2}E_2^2E_3[N_2\mathfrak{D}(\omega-\nu_{12})+\bar{N}\mathfrak{D}(\Delta)][\mathfrak{D}_a(\Delta)+\mathfrak{D}_b(\Delta)]$$
$$\times\exp-i((2\nu_2-\nu_3)t-(2\varphi_2-\varphi_3))]+\text{complex conjugate.} \quad (104)$$

The in-phase coefficient $C_1^{(3)}$ is then

$$C_1^{(3)}=\tfrac{1}{8}\pi^{1/2}[\wp^4/(\hbar^3Ku)][E_1^3\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b))(\omega-\nu_1)\mathfrak{L}(\omega-\nu_1)+E_1E_2^2\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b))[(\omega-\nu_{12})\mathfrak{L}(\omega-\nu_{12})+(2/\Delta)]$$
$$+E_1E_3^2\bar{N}(\gamma_{ab}/(\gamma_a\gamma_b)[(\omega-\nu_2)\mathfrak{L}(\omega-\nu_2)+(1/\Delta)]+E_2^2E_1N_2\gamma_{ab}(\omega-\nu_1+\Delta)\mathfrak{L}(\omega-\nu_1)/\Delta^2$$
$$+\tfrac{1}{4}E_3^2E_1N_4\gamma_{ab}(\omega-\nu_1+2\Delta)\mathfrak{L}(\omega-\nu_1)/\Delta^2+E_2^2E_3\Delta^{-2}\{N_2\gamma_{ab}(\omega-\nu_{12}+\Delta)\mathfrak{L}(\omega-\nu_{12})\cos\psi$$
$$+[N_2(\gamma_{ab}^2-(\omega-\nu_{12})\Delta)\mathfrak{L}(\omega-\nu_{12})-\bar{N}]\sin\psi\}], \quad (105)$$

where the "relative phase angle" ψ is defined as

$$\psi=(2\nu_2-\nu_1-\nu_3)t+(2\varphi_2-\varphi_1-\varphi_3). \quad (106)$$

The quadrature coefficient $S_1^{(3)}$ is given by

$$S_1^{(3)}=\tfrac{1}{8}\pi^{1/2}[\wp^4/(\hbar^3Ku)]\{E_1^3\bar{N}[\gamma_{ab}^2\mathfrak{L}(\omega-\nu_1)+1]/(\gamma_a\gamma_b)+E_1E_2^2\bar{N}[\gamma_{ab}^2\mathfrak{L}(\omega-\nu_{12})+4\gamma_{ab}^2\Delta^{-2}]/(\gamma_a\gamma_b)$$
$$+E_1E_3^2\bar{N}[\gamma_{ab}^2\mathfrak{L}(\omega-\nu_2)+\gamma_{ab}^2\Delta^{-2}]/(\gamma_a\gamma_b)+E_2^2E_1\Delta^{-2}[N_2(\gamma_{ab}^2-(\omega-\nu_1)\Delta)\mathfrak{L}(\omega-\nu_1)-2\bar{N}]$$
$$+\tfrac{1}{4}E_3^2E_1\Delta^{-2}[N_4(\gamma_{ab}^2-2(\omega-\nu_1)\Delta)\mathfrak{L}(\omega-\nu_1)-2\bar{N}]+E_2^2E_3\Delta^{-2}[[N_2(\gamma_{ab}^2-(\omega-\nu_{12})\Delta)\mathfrak{L}(\omega-\nu_{12})-\bar{N}]\cos\psi$$
$$-N_2\gamma_{ab}(\omega-\nu_{12}+\Delta)\mathfrak{L}(\omega-\nu_{12})\sin\psi]\}. \quad (107)$$

There are similar expressions for the other coefficients which appear in the self-consistent field equations (13) and (14).

The generalization of Eq. (81) takes the form

$$\dot{E}_1=\alpha_1E_1-\beta_1E_1^3-\theta_{12}E_1E_2^2-\theta_{13}E_1E_3^2-(\eta_{23}\cos\psi+\xi_{23}\sin\psi)E_2^2E_3,$$
$$\dot{E}_2=\alpha_2E_2-\beta_2E_2^3-\theta_{21}E_2E_1^2-\theta_{23}E_2E_3^2-(\eta_{13}\cos\psi+\xi_{23}\sin\psi)E_1E_2E_3, \quad (108)$$
$$\dot{E}_3=\alpha_3E_3-\beta_3E_3^3-\theta_{31}E_3E_1^2-\theta_{32}E_3E_2^2-(\eta_{21}\cos\psi+\xi_{21}\sin\psi)E_2^2E_1.$$

The coefficients α_n and β_n were already calculated in the single-frequency case, and are given by Eqs. (82) and (83). The other coefficients are given by

$$\theta_{12}=\tfrac{1}{16}\pi^{1/2}\nu(\wp^4/(\epsilon_0\hbar^3Ku))\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_{12})+4\Delta^{-2}]-2\bar{N}\Delta^{-2}+N_2\Delta^{-2}[\gamma_{ab}^2-(\omega-\nu_1)\Delta]\mathfrak{L}(\omega-\nu_1)\}, \quad (109)$$

$$\theta_{13}=\tfrac{1}{16}\pi^{1/2}\nu(\wp^4/(\epsilon_0\hbar^3Ku))\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_2)+\Delta^{-2}]-\tfrac{1}{2}\bar{N}\Delta^{-2}+\tfrac{1}{4}N_4\Delta^{-2}[\gamma_{ab}^2-2(\omega-\nu_1)\Delta]\mathfrak{L}(\omega-\nu_1)\}, \quad (110)$$

$$\eta_{23}=\tfrac{1}{16}\pi^{1/2}\nu(\wp^4/(\epsilon_0\hbar^3Ku\Delta^2))\{N_2[\gamma_{ab}^2-(\omega-\nu_{12})\Delta]\mathfrak{L}(\omega-\nu_{12})-\bar{N}\}, \quad (111)$$

$$\xi_{23}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}N_2\gamma_{ab}(\omega-\nu_{12}+\Delta)\mathfrak{L}(\omega-\nu_{12}), \quad (112)$$

$$\theta_{21}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_{12})+4\Delta^{-2}]+N_2\Delta^{-2}[\gamma_{ab}^2+(\omega-\nu_2)\Delta]\mathfrak{L}(\omega-\nu_2)-2\bar{N}\Delta^{-2}\}, \quad (113)$$

$$\theta_{23}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_{23})+4\Delta^{-2}]+N_2\Delta^{-2}[\gamma_{ab}^2-(\omega-\nu_2)\Delta]\mathfrak{L}(\omega-\nu_2)-2\bar{N}\Delta^{-2}\}, \quad (114)$$

$$\xi_{13}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}N_2\gamma_{ab}[(\omega-\nu_{23}-\Delta)\mathfrak{L}(\omega-\nu_{23})+(\omega-\nu_{12}+\Delta)\mathfrak{L}(\omega-\nu_{12})], \quad (115)$$

$$\eta_{13}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}\{N_2[\gamma_{ab}^2+(\omega-\nu_{23})\Delta]\mathfrak{L}(\omega-\nu_{23})+N_2[\gamma_{ab}^2-(\omega-\nu_{12})\Delta]\mathfrak{L}(\omega-\nu_{12})+2\bar{N}\}, \quad (116)$$

$$\theta_{31}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_2)+\Delta^{-2}]+\tfrac{1}{4}N_4\Delta^{-2}[\gamma_{ab}^2+2(\omega-\nu_3)\Delta]\mathfrak{L}(\omega-\nu_3)-\tfrac{1}{2}\bar{N}\Delta^{-2}\}, \quad (117)$$

$$\theta_{32}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}\gamma_{ab}^2(\gamma_a\gamma_b)^{-1}[\mathfrak{L}(\omega-\nu_{23})+4\Delta^{-2}]+N_2\Delta^{-2}[\gamma_{ab}^2+(\omega-\nu_3)\Delta]\mathfrak{L}(\omega-\nu_3)-2\bar{N}\Delta^{-2}\}, \quad (118)$$

$$\eta_{21}=\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}\{N_2[\gamma_{ab}^2+(\omega-\nu_{23})\Delta]\mathfrak{L}(\omega-\nu_{23})-\bar{N}\}, \quad (119)$$

$$\xi_{21}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}N_2\gamma_{ab}(\omega-\nu_{23}-\Delta)\mathfrak{L}(\omega-\nu_{23}). \quad (120)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The frequency and phase determining equations are of the form

$$\nu_1+\dot{\varphi}_1=\Omega_1+\sigma_1+\rho_1E_1^2+\tau_{12}E_2^2+\tau_{13}E_3^2+E_2^2E_3E_1^{-1}(\eta_{23}\sin\psi-\xi_{23}\cos\psi), \quad (121)$$

$$\nu_2+\dot{\varphi}_2=\Omega_2+\sigma_2+\rho_2E_2^2+\tau_{21}E_1^2+\tau_{23}E_3^2+E_1E_3(\eta_{13}\sin\psi-\xi_{13}\cos\psi), \quad (122)$$

$$\nu_3+\dot{\varphi}_3=\Omega_3+\sigma_3+\rho_3E_3^2+\tau_{31}E_1^2+\tau_{32}E_2^2+E_2^2E_1E_3^{-1}(\eta_{21}\sin\psi-\xi_{21}\cos\psi), \quad (123)$$

where the coefficients σ_n and ρ_n are already given by Eqs. (90) and (91), and the η's and ξ's by Eqs. (111), (112), (115), (116), (119), and (120). The remaining coefficients are

$$\tau_{12}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_{12})\mathfrak{L}(\omega-\nu_{12})+2\gamma_{ab}\Delta^{-1}]+N_2\Delta^{-2}\gamma_{ab}(\omega-\nu_1+\Delta)\mathfrak{L}(\omega-\nu_1)\}, \quad (124)$$

$$\tau_{13}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_2)\mathfrak{L}(\omega-\nu_2)+\gamma_{ab}\Delta^{-1}]+\tfrac{1}{4}N_4\Delta^{-2}\gamma_{ab}(\omega-\nu_1+2\Delta)\mathfrak{L}(\omega-\nu_1)\}, \quad (125)$$

$$\tau_{21}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_{12})\mathfrak{L}(\omega-\nu_{12})-2\gamma_{ab}\Delta^{-1}]+N_2\Delta^{-2}\gamma_{ab}(\omega-\nu_2-\Delta)\mathfrak{L}(\omega-\nu_2)\}, \quad (126)$$

$$\tau_{23}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_{23})\mathfrak{L}(\omega-\nu_{23})+2\gamma_{ab}\Delta^{-1}]+N_2\Delta^{-2}\gamma_{ab}(\omega-\nu_2+\Delta)\mathfrak{L}(\omega-\nu_2)\}, \quad (127)$$

$$\tau_{31}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_2)\mathfrak{L}(\omega-\nu_2)-\gamma_{ab}\Delta^{-1}]+\tfrac{1}{4}N_4\Delta^{-2}\gamma_{ab}(\omega-\nu_3-2\Delta)\mathfrak{L}(\omega-\nu_3)\}, \quad (128)$$

$$\tau_{32}=-\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku)^{-1}\{\bar{N}(\gamma_a\gamma_b)^{-1}[\gamma_{ab}(\omega-\nu_{23})\mathfrak{L}(\omega-\nu_{23})-2\gamma_{ab}\Delta^{-1}]+N_2\Delta^{-2}\gamma_{ab}(\omega-\nu_3-\Delta)\mathfrak{L}(\omega-\nu_3)\}. \quad (129)$$

10. TWO-FREQUENCY OPERATION

We may here drop all terms referring to the third frequency ν_3. There are now no "combination tones" in near resonance with Ω_1 and Ω_2. The amplitudes E_1 and E_2 are determined by the differential equations

$$\dot{E}_1=\alpha_1E_1-\beta_1E_1^3-\theta_{12}E_1E_2^2,$$
$$\dot{E}_2=\alpha_2E_2-\theta_{21}E_2E_1^2-\beta_2E_2^3, \quad (130)$$

where the coefficients α_n, β_n, θ_{12}, and θ_{21} are now given by Eqs. (82), (83), (109), and (113).

Introducing the squared amplitudes

$$X=E_1^2 \quad \text{and} \quad Y=E_2^2, \quad (131)$$

Eqs. (126) become

$$\dot{X}=2X(\alpha_1-\beta_1X-\theta_{12}Y),$$
$$\dot{Y}=2Y(\alpha_2-\theta_{21}X-\beta_2Y). \quad (132)$$

The condition for a steady state of oscillation is $\dot{X}=0$, $\dot{Y}=0$ and may represent graphically in an X-Y plane by the point of intersection of the two straight lines

$$L_1: \quad \beta_1X+\theta_{12}Y=\alpha_1,$$
$$L_2: \quad \theta_{21}X+\beta_2Y=\alpha_2, \quad (133)$$

if there is one in the first quadrant, together with the single-frequency solutions

$$X=\alpha_1/\beta_1, \quad Y=0 \quad \text{and} \quad X=0, \quad Y=\alpha_2/\beta_2. \quad (134)$$

The differential equations (132) allow us to follow the temporal behavior of the state X, Y of oscillation in the phase plane X, Y. Through any point in this plane (except stationary points) there passes a curve which indicates the path followed by the representative point (X,Y) on its way to a stable state of oscillation. The parametric equations of the curves are $X=X(t)$, $Y=Y(t)$ with the time t as parameter. From the differential equations it is seen that each curve has a vertical tangent when it crosses the first of the straight lines (133), and a horizontal tangent when it crosses the second straight line. This principle facilitates a very simple, if qualitative, graphical integration of the differential equations (132) for the paths of the phase points in the cases discussed below.

The various possibilities are depicted[18] in Figs. 3–5 where it is assumed that both α_1 and α_2 are positive,

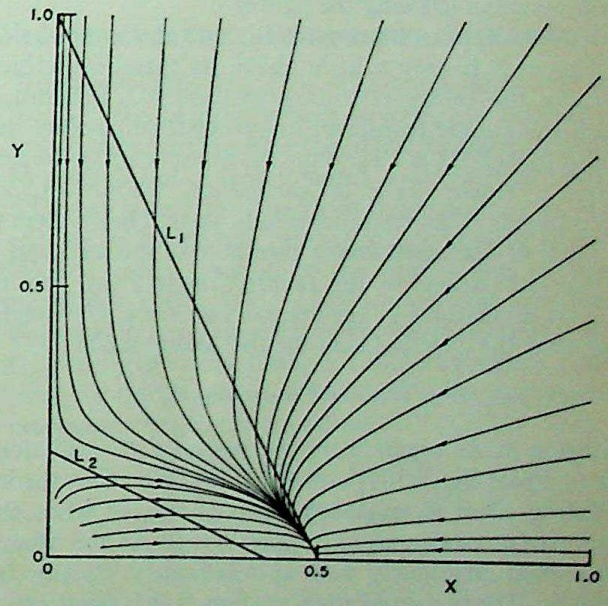


FIG. 3. Phase curves showing the transient behavior of two-mode oscillation. The straight lines L_1 and L_2 of Eq. (133) are taken to have coefficients $\alpha_1=1$, $\alpha_2=0.4$, $\beta_1=\beta_2=2$, $\theta_{12}=\theta_{21}=1$. The slope of a phase curve is zero when it crosses line L_2 and infinite when it crosses line L_1. Although both modes are above threshold, the favored X oscillation is able to quench the Y oscillation.

[18] The phase paths of Figs. 2-4 were kindly integrated on an analog computer by Dr. B. Wise of the Engineering Science Laboratory, Oxford, to whom the author is very indebted.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

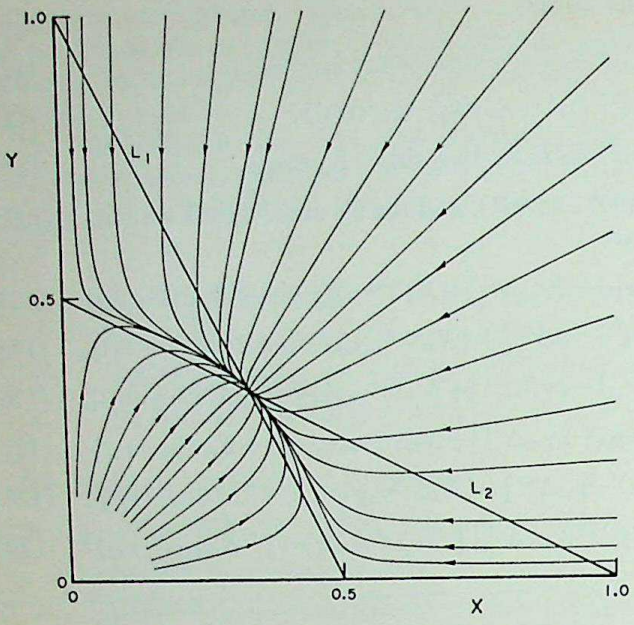


FIG. 4. Diagram similar to Fig. 3, except that the gain parameter for the second mode has been raised to $\alpha_2=1$. Simultaneous oscillations at both frequencies occur at the single stable steady state. Both Figs. 3 and 4 correspond to "weak" coupling.

so that the two modes are individually above threshold. The coefficients β_n are necessarily positive if the optical medium is an active one, and while the mode competition coefficients θ could conceivably have the opposite sign, they have been assumed positive (and equal to each other) in drawing the figures.

Figure 3 applies when mode 1 is well above threshold, but mode 2 is only a little above its threshold, either because the cavity resonance frequency Ω_2 is detuned from the atomic transition frequency ω or possibly because $Q_2<Q_1$. It is clear that the point $(\alpha_1/\beta_1,0)$ represents a stable state of oscillation, while $(0,\alpha_2/\beta_2)$ corresponds to an unstable steady state. Hence there is a range of operation above threshold of modes 1 and 2 where oscillations in the favored mode 1 are able to inhibit oscillations at the second frequency. One might say that the effective gain for the second mode

$$\alpha_2'=\alpha_2-\theta X=\alpha_2-\theta\alpha_1/\beta_1$$

is being made negative by the presence of oscillations at ν_1. From Eqs. (109) and (113) it is seen that the inhibiting effect is enhanced when $\nu_{12}=\omega$, or when the two cavity modes are on opposite sides of the atomic transition frequency, and approximately equally far from it. The physical interpretation of this effect which clearly involves o.f. saturation will be brought out more clearly in Sec. 18.

As the excitation increases, α_2' will eventually become positive, and the relevant diagrams are Figs. 4 and 5. The former applies when $\beta_1\beta_2>\theta^2$ (weak coupling) and the latter when $\beta_1\beta_2<\theta^2$ (strong coupling).

The two cases of weak and strong coupling give very different behaviors. For weak coupling the point of intersection of the two straight lines gives stable steady-state operation, while the single-frequency operating points are unstable. The optical maser oscillates simultaneously at two frequencies under these conditions. For strong coupling, on the other hand, the point of intersection of the two straight lines represents an unstable steady state and would not be realized in practice. All other points in the state diagram evolve into one or the other of the single-frequency operation points. Which of the two is reached depends on the past history of the state of oscillation. In other words, there is hysteresis.

In the Doppler broadened gaseous optical maser Eqs. (83), (109), and (113) indicate that the case of weak coupling is naturally favored, since $\beta_1\beta_2$ tends to be greater than $\theta_{12}\theta_{21}$. Hence, with possible exceptions such as the one discussed in the next section, double frequency operation is preferred. However, when van der Pol's theory of a double resonance feed back triode oscillator is transcribed into our notation, one finds that in his case $\theta=2\beta$. This results from his assumption of a term $(E_1\cos\nu_1 t+E_2\cos\nu_2 t)^3$ in the triode output current. After discarding terms which have frequencies far from ν_1 and ν_2, this becomes

$$\tfrac{3}{4}(E_1^3+2E_1E_2^2)\cos\nu_1 t+\tfrac{3}{4}(E_2^3+2E_2E_1^2)\cos\nu_2 t,$$

which leads to the stated relation $\theta=2\beta$. The van der Pol oscillator prefers operation at a single frequency and exhibits hysteresis phenomena, as in the case of strong coupling. Evidently the atomic medium of an optical

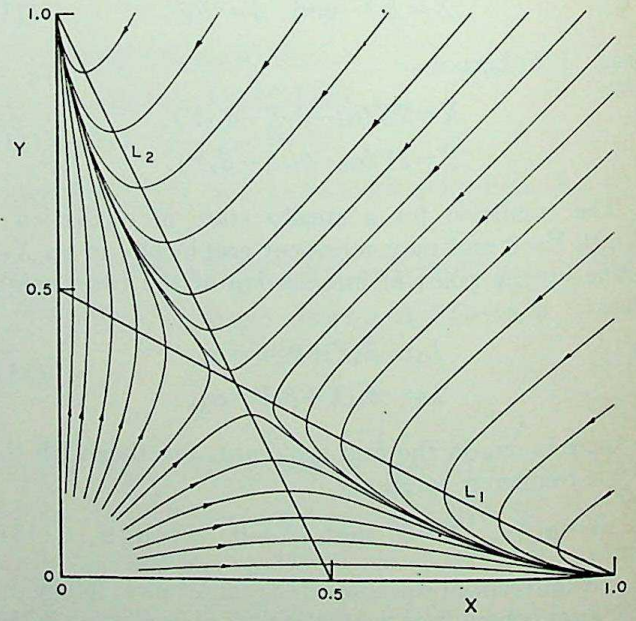


FIG. 5. Phase curves showing the transient behavior of two-mode oscillation when the straight lines L_1 and L_2 of Eq. (133) are taken to have the coefficients $\alpha_1=\alpha_2=1$, $\beta_1=\beta_2=1$, $\theta_{12}=\theta_{21}=2$ (strong coupling). There are two possible stable steady states, each corresponding to single-frequency operation. The particular state reached depends on the initial conditions. Hysteresis phenomena would occur if the parameters characterizing the oscillator were slowly changed.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

maser differs from van der Pol's triode oscillator because the nonlinear response of the atoms to frequencies ν_1 and ν_2 has a resonant character not assumed of the triode. It will become clearer from the discussion of Sec. 18 that two groups of atoms with different velocity are driving the two oscillations with only a limited degree of interference. One could easily make a comparable model of the van der Pol type with *two* triodes each with its own tank circuit. A small amount of coupling between the two oscillators would produce intermode effects described by the θ coefficients in the weak-coupling case.

There is also a further case in which the two straight lines coincide, i.e., when $\alpha_1/\alpha_2=\beta_1/\theta=\theta/\beta_2$. In that case, there is a neutral steady state for the representative point lying anywhere on the line within the first quadrant. In practice, the state of operation when this condition is nearly satisfied should be very sensitive to microphonic disturbances.

11. INTENSITIES AND FREQUENCIES IN TWO-MODE OPERATION

The two-mode steady-state solution of (130) is given by

$$\begin{aligned} E_1^2&=(\beta_2\alpha_1-\theta_{12}\alpha_2)/(\beta_1\beta_2-\theta_{12}\theta_{21}),\\ E_2^2&=(\beta_1\alpha_2-\theta_{21}\alpha_1)/(\beta_1\beta_2-\theta_{12}\theta_{21}),\end{aligned} \tag{135}$$

where, as explained before, the right-hand sides may be evaluated for $\nu_1=\Omega_1$ and $\nu_2=\Omega_2$ without appreciable error. The frequencies are obtained by dropping inapplicable terms from Eqs. (121)–(123). We may set $\dot\varphi_1=\dot\varphi_2$ without loss of generality, and find

$$\begin{aligned} \nu_1&=\Omega_1+\sigma_1+\rho_1E_1^2+\tau_{12}E_2^2,\\ \nu_2&=\Omega_2+\sigma_2+\rho_2E_2^2+\tau_{21}E_1^2,\end{aligned} \tag{136}$$

where the right-hand sides are to be evaluated for $\nu_n=\Omega_n$, and the E_n^2 are as given in (135).

These equations are fairly complicated, and probably can be used in full generality only for a numerical analysis[19] of very detailed data on optical maser operation. Such a study would be simplified if the values of the E's could be inferred experimentally, since that would effectively reduce the dependence of Eqs. (136) on cavity tuning. For the present we will merely work out the frequencies for the important case of "midtuning" where $\nu_{12}=\frac{1}{2}(\nu_1+\nu_2)=\omega$, $\omega-\nu_1=\frac{1}{2}\Delta$ and $\nu_2-\omega=\frac{1}{2}\Delta$. If we regard Δ as being much greater than the γ's, we then may approximate the coefficients appearing in Eqs. (135)–(136) as follows:

$$\beta_1\simeq\beta_2\simeq\theta_{12}\simeq\theta_{21}\simeq\tfrac{1}{16}\pi^{1/2}\wp^4\bar N(\epsilon_0\hbar^3Ku\gamma_a\gamma_b)^{-1}=\beta, \tag{137}$$

$$\rho_2\approx-\rho_1=2\gamma_{ab}\Delta^{-1}\beta=\rho, \tag{138}$$

$$\tau_{12}\approx-\tau_{21}\approx-\rho, \tag{139}$$

and assuming $Q_1=Q_2=Q$, the gain parameters become

$$\alpha_1\approx\alpha_2\simeq\tfrac{1}{2}\nu Q_n^{-1}\{\Re[Z_i(\tfrac{1}{2}\Delta)/Z_i(0)]-1\}=\alpha. \tag{140}$$

The intensities are $E_1^2=E_2^2=\frac{1}{2}\alpha/\beta$, and the frequencies are

$$\begin{aligned} \nu_1&=\Omega_1+\sigma_1-\rho\alpha/\beta,\\ \nu_2&=\Omega_2+\sigma_2+\rho\alpha/\beta,\end{aligned} \tag{141}$$

so that the beat frequency is

$$\nu_2-\nu_1\simeq\Delta+(\sigma_2-\sigma_1)+4\gamma_{ab}\alpha/\Delta. \tag{142}$$

It will be seen that in the approximations of Eq. (137) we have $\beta_1\beta_2=\theta_{12}\theta_{21}$, and hence in order to decide whether the coupling is "weak" or "strong" it is necessary to evaluate the coefficients β and θ more exactly. In case of exact midtuning, one finds that $\beta^2-\theta^2$ has the same sign as $\bar N+N_2$. Ordinarily this would be positive since $\bar N$ must be positive for oscillations to occur. However, it is possible in principle to arrange to have $\bar N>0$ and $N_2<-\bar N$ by having $N(z)>0$ in the middle two quarters of the tube length, and $N(z)<0$ in the end quarters. In practice the last requirement could be met by adjusting the gas discharge conditions near the ends so that the lower maser atomic level is more populated than the upper one. In He-Ne masers an increase of tube diameter near the ends might be helpful in this respect. Diffusion of Ne metastables to the walls is thereby reduced, and electron excitation of the lower maser level is increased.

12. NORMAL THREE-FREQUENCY OPERATION

Equations (108) and (121)–(123) are fairly complicated, but can be readily used to discuss a number of special cases, as indicated in the following three sections.

In general, unless care is taken to adjust the cavity tuning very accurately, the three frequencies ν_1, ν_2, and ν_3 will be such that the relative phase angle ψ of Eq. (106) is a linear function of the time. Then the last terms in Eqs. (108), (121)–(123) are periodic functions of time, and in some approximation their effects average out. If we neglect these terms, we can get a steady-state solution for the intensities E_1^2, E_2^2, E_3^2 from the system of inhomogeneous linear equations

$$\begin{aligned} \alpha_1&=\beta_1E_1^2+\theta_{12}E_2^2+\theta_{13}E_3^2,\\ \alpha_2&=\beta_2E_2^2+\theta_{21}E_1^2+\theta_{23}E_3^2,\\ \alpha_3&=\beta_3E_3^2+\theta_{31}E_1^2+\theta_{32}E_2^2,\end{aligned} \tag{143}$$

and for the frequencies ν_1, ν_2, ν_3 from

$$\begin{aligned} \nu_1&=\Omega_1+\sigma_1+\rho_1E_1^2+\tau_{12}E_2^2+\tau_{13}E_3^2,\\ \nu_2&=\Omega_2+\sigma_2+\rho_2E_2^2+\tau_{21}E_1^2+\tau_{23}E_3^2,\\ \nu_3&=\Omega_3+\sigma_3+\rho_3E_3^2+\tau_{31}E_1^2+\tau_{32}E_2^2,\end{aligned} \tag{144}$$

again taking $\dot\varphi_1=\dot\varphi_2=\dot\varphi_3=0$. Since the coefficients in Eqs. (143) and (144) are slowly varying functions of frequency, it will suffice first to determine the E_n^2 from

[19] Such an analysis is being made by Dr. R. L. Fork and Dr. M. A. Pollack. The author is very grateful to them for helpful discussions on this and other parts of the manuscript.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(143) and then to calculate the ν_n's from (144). The equations are a fairly obvious generalization of those for the two frequency case. It will be noted that operation with $E_1^2 \neq 0$, $E_2^2 \neq 0$ can inhibit normal oscillation at ν_3 until, with increasing excitation,

$$\alpha_3' = \alpha_3 - \theta_{31}E_1^2 - \theta_{32}E_2^2 \tag{145}$$

becomes positive.

Further discussion of pulling, pushing, and mode competition for this case will be left to readers requiring a numerical analysis of their data.

13. COMBINATION TONES

As mentioned in Secs. 7, 9, the third-order polarization $P_n^{(3)}(t)$ of the active medium has constituents which oscillate at all possible frequencies of the form $\nu_\mu - \nu_\rho + \nu_\sigma$. Even for "two"-frequency oscillation, there are additional frequencies $2\nu_2 - \nu_1 \equiv \nu_3'$ and $2\nu_1 - \nu_2 \equiv \nu_0'$ in the polarization which are very close to resonance with the principal cavity modes just above and below the two Ω_1 and Ω_2 of main interest. As a consequence of Maxwell's equations, fields at frequencies ν_3' and ν_0' necessarily exist in the cavity and can appear in the output. Well below the threshold for normal three-frequency oscillation, E_3 (and also E_0) will be much smaller than E_1 and E_2. We also assume $\gamma_{ab} \ll \Delta \ll Ku$ and $\nu_2 \simeq \omega$ so that

$$\eta_{21} \simeq -\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3 Ku\Delta^2)^{-1}[\bar{N}+2N_2], \tag{146}$$
$$\xi_{21} \simeq 0.$$

For a steady state, the third Eq. (108) then gives

$$E_3 \simeq \eta_{21}E_2^2E_1[\alpha_3 - \theta_{31}E_1^2 - \theta_{32}E_2^2]^{-1}\cos\psi. \tag{147}$$

The relative phase angle ψ is determined by Eq. (123) to have $\sin\psi \simeq 0$, whence $|\cos\psi| \simeq 1$, and

$$E_3 \approx -\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3 Ku\Delta^2)^{-1} \times[\alpha_3 - \theta_{31}E_1^2 - \theta_{32}E_2^2]^{-1}|\bar{N}+2N_2|E_2^2E_1. \tag{148}$$

This expression is intended to be used under excitation conditions for which the denominator is negative. When the excitation increases, and the factor involving α_3 turns positive, the neglected nonlinear terms in E_3 would have to be taken into account in order to describe the previously discussed normal three-frequency operation.

In order to obtain the combination tone $(2\nu_2 - \nu_1)$ experimentally, one should adjust the cavity tuning so Ω_2 is slightly above the atomic transition frequency ω, thereby making Ω_1 a little nearer resonance than Ω_3. The excitation should be increased until "two"-frequency operation is obtained, but not yet genuine three-frequency operation.

Under these conditions, the θ's in Eq. (148) contain a factor Δ^{-2} and to simplify the discussion they will now be neglected. Equation (148) then has a factor in its denominator

$$\alpha_3 \approx -\tfrac{1}{2}(\nu/Q_3)/G_3, \tag{149}$$

where the gain factor G_3 is given approximately by

$$G_3 \approx [1-(\bar{N}/\bar{N}_T)(Z_i(\Delta)/Z_i(0))]^{-1}, \tag{150}$$

if one is not too near to threshold for normal three-frequency operation. Equations (148)–(150), within their domain of validity, indicate that E_3 is smaller than E_1 by a factor

$$E_3/E_1 \approx \tfrac{1}{8}\pi^{1/2}\wp^4E_2^2(\epsilon_0\hbar^3 Ku\Delta^2)^{-1}Q_3G_3|\bar{N}+2N_2|. \tag{151}$$

Using Eq. (62) for $n=3$, this can be expressed in a convenient form

$$E_3/E_1 = \tfrac{1}{4}[\tfrac{1}{2}(\wp E_2)^2/(\hbar^2\gamma_a\gamma_b)][(\gamma_a\gamma_b)/\Delta^2]G_3 \times|1+(2N_2/\bar{N})|\exp(\Delta/Ku)^2, \tag{152}$$

which shows how the amplitude of the combination tone depends on G_3, and on the saturation parameter as given by Eq. (96) for $n=2$. It should be noted that Eq. (152) could vanish if the spatial distribution of the excitation density is such that $2N_2+\bar{N}=0$. (If the excitation is confined to the central region of the Fabry-Perot tube, N_2 and $\bar{N}$, by Eqs. (47), (58) have opposite signs.) An experimental study of the above phenomena might facilitate determination of some of the quantities which enter into our equations but for which direct experimental values are not yet available.

14. FREQUENCY LOCKING PHENOMENA

It has been observed by Javan[20] and by Fork[19] that when the cavity tuning is gradually changed in normal three-frequency operation so that the separation of the beat notes $\nu_2 - \nu_1$ and $\nu_3 - \nu_2$ approaches a small value (typically of order 1 kc/sec), a frequency jump occurs. This phenomenon can be easily understood by reference to Eqs. (121)–(123). For simplicity, we neglect the small frequency pushing associated with the terms involving ρ_{nm} and τ_{nm}, since the nonlinear pulling terms σ_n already give sufficient generality to the frequency relationships. By subtracting the sum of Eqs. (121)–(123) from twice Eq. (122), we find a differential equation for the relative phase angle ψ of Eq. (106) in the form

$$\dot{\psi} = \sigma + A\sin\psi + B\cos\psi, \tag{153}$$

where

$$\sigma = 2\sigma_2 - \sigma_1 - \sigma_3, \tag{154}$$

and A and B are slowly varying quantities which depend on the E_n, ξ_{nm}, and η_{nm}. We evaluate the ξ_{nm} and η_{nm} with the usual approximations $\gamma_{ab} \ll \Delta \ll Ku$. Then

$$\eta_{23} \simeq \eta_{21} \simeq -\tfrac{1}{16}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3 Ku\Delta^2)^{-1}\{\bar{N}+2N_2\},$$
$$\eta_{13} \simeq \tfrac{1}{8}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3 Ku\Delta^2)\{\bar{N}-2N_2\}, \tag{155}$$
$$\xi_{23} \simeq \xi_{21} \simeq \xi_{13} \simeq 0.$$

Because of the nearly symmetrical tuning of ν_1 and ν_3

[20] A. Javan, private communications for which the author is very grateful.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

we may set $E_1=E_3$ and find

$$A\approx\tfrac{1}{8}\pi^{1/2}\nu\wp^4(\epsilon_0\hbar^3Ku\Delta^2)^{-1}\times\{(\bar{N}+2N_2)E_2^2+2(\bar{N}-2N_2)E_1^2\}, \quad (156)$$

$B\approx 0$.

The differential equation

$$\dot{\psi}=\sigma+A\sin\psi+B\cos\psi \quad (157)$$

has an implicit solution

$$t(\psi)=\int_{\psi_0}^{\psi}dx/(\sigma+A\sin x+B\cos x), \quad (158)$$

where ψ_0 is the value of ψ at $t=0$. The character of the function $\psi(t)$ depends critically on whether

$$(A^2+B^2)^{1/2}<|\sigma| \quad \text{or} \quad (A^2+B^2)^{1/2}>|\sigma|.$$

In the first case, the integrand of (158) has no singularities in the range of integration, and one finds that as ψ approaches infinity so does t. Asymptotically, apart from pulsations, one has $\psi\simeq\sigma t+\text{const}$. On insertion of this value in Eqs. (121)–(123) one finds that the frequencies are given by

$$\nu_n\sim\Omega_n+\sigma_n \quad (159)$$

apart from pushing effects and pulsations of phase φ_n, in agreement with the results of Sec. 6 with the approximations made here.

In the second case, $(A^2+B^2)>\sigma^2$, the integral diverges, i.e., $t\to\infty$ when ψ reaches the value $-\sin^{-1}\times(\sigma/(A^2+B^2)^{1/2}$. In other words, $\psi(t)$ approaches this value asymptotically. The disappearance from ψ of any linear dependence on t forces $\nu_3-\nu_2=\nu_2-\nu_1$ and frequency locking ensues, with a definite relative phase angle $2\varphi_2-\varphi_1-\varphi_3$.

Let us suppose that the maser is in normal three-frequency operation with two distinct beat notes $\nu_3-\nu_2$ and $\nu_2-\nu_1$ near to $\Delta\approx 150$ Mc/sec, i.e., $\sigma=(\nu_3-\nu_2)-(\nu_2-\nu_1)$ is somewhat greater than $(A^2+B^2)^{1/2}$. As the middle cavity frequency Ω_2 is tuned closer to the atomic resonance frequency ω, the separation of beat note frequencies $|\sigma|$ decreases. There should be some pulsations in phase which would increase in amplitude as symmetrical tuning is approached. When $|\sigma|$ reaches $(A^2+B^2)^{1/2}$, a quick transition to the locked state should be made, and only one beat note should be observed. Under the additional simplifying assumption $E_2\gg E_1\approx E_3$, and with use of the starting condition (60) for single-frequency oscillation, the separation of the two beat notes which could be attained just before locking occurs should be given by

$$|\sigma|=\tfrac{1}{8}[(\bar{N}+2N_2)/\bar{N}_T](\wp E_2/\hbar\Delta)^2\nu/Q, \quad (160)$$

which is conveniently expressed as a small fraction of the cavity bandwidth.

It might be pointed out that the above phenomenon is very closely related to one discussed by van der Pol[21] in 1924–27. He considered a self-sustained triode oscillator, capable of oscillation at frequency ν_1. If an external signal at ν is injected into the tank circuit, it may be possible to detect a beat note at $|\nu-\nu_1|$ using a square-law detector. If, however, ν is tuned gradually towards ν_1, a very sudden jump occurs, after which oscillations occur only at ν and the beat note disappears. The width of the "quiet" frequency range depends approximately linearly on the amplitude of the injected signal, when this is small. In the case of the optical maser, we can think that an oscillator at ν_1 is being perturbed by an "external" signal at the combination tone frequency $\nu=2\nu_2-\nu_3$ which arises from the third-order polarization $P_1^{(3)}(t)$ induced in the nonlinear active medium.

15. POPULATION CHANGES AND PULSATIONS

In the absence of o.f. oscillations, the density of atoms in one of the two maser states, say a, can be determined by suitable integrations from $\rho_{aa}^{(0)}(a,z_0,t_0,v,t)$ as given by Eq. (24). When oscillations set in, there are contributions of second order which can be calculated from $\rho_{aa}^{(2)}(\alpha,z_0,t_0,v,t)$ using Eq. (31) for $\alpha=a$ and a similar equation for $\alpha=b$. One has

$$\rho_{aa}(z,t)=\int_{-\infty}^{\infty}dv\int_{-\infty}^{t}dt_0\int dz_0\delta(z-z_0-vt+vt_0)\times\sum_{\alpha=a,b}\lambda_\alpha(z_0,t_0,v)\rho_{aa}(\alpha,z_0,t_0,v,t). \quad (161)$$

It will suffice merely to give the result. One finds, with obvious approximations

$$\rho_{aa}(z,t)=[\Lambda_a(z,t)/\gamma_a]+\{[\Lambda_a(z,t)/\gamma_a]-[\Lambda_b(z,t)/\gamma_b]\}\times\tfrac{1}{8}(\wp^2/Ku)\sum_\mu\sum_\rho\{E_\mu E_\rho\mathfrak{D}_a(\nu_\mu-\nu_\rho)iZ(\omega-\nu_\rho)\times\exp i[(\nu_\mu-\nu_\rho)t+(\varphi_\mu-\varphi_\rho)]+\text{c.c.}\}\times\cos[(n_\mu-n_\rho)\pi z/L] \quad (162)$$

apart from terms with rapid spatial oscillations.

For single-frequency operation, one finds

$$\rho_{aa}(z,t)=[\Lambda_a(z,t)/\gamma_a]-\{[\Lambda_a(z,t)/\gamma_a]-[\Lambda_b(z,t)/\gamma_b]\}\times\tfrac{1}{4}(\wp E_1)^2(\gamma_aKu)^{-1}Z_i(\omega-\nu_1), \quad (163)$$

which contains the lowest order effects of o.f. saturation. In some cases, the density could be monitored by observation of the decay radiation emitted from state a (or b) in transition to some lower level. (Of course, if trapping of resonance radiation were involved, the interpretation would be somewhat complicated.) The change produced by o.f. radiation in $\rho_{aa}(z,t)$ might serve to aid in the determination of parameters like

[21] B. van der Pol, Phil. Mag. 3, 65 (1927) and the review article cited in Ref. 17.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

γ_a, E_1^2, etc., which enter into all our equations, but which might not otherwise be known.

For two-frequency operation of the optical maser, ρ_{aa} has a pulsating constituent at a frequency near Δ besides a dc part. If we assume $\gamma_{ab}\sim\gamma_a\gg\Delta\ll Ku$

$$\begin{aligned}\rho_{aa}(z,t)&=[\Lambda_a(z,t)/\gamma_a]-\{[\Lambda_a(z,t)/\gamma_a]-[\Lambda_b(z,t)/\gamma_b]\}\\&\times\tfrac{1}{4}\pi^{1/2}\wp^2(Ku\gamma_a)^{-1}[E_1^2\exp-(\omega-\nu_1)^2/(Ku)^2\\&+E_2^2\exp-(\omega-\nu_2)^2/(Ku)^2]+\{[\Lambda_a/\gamma_a]-[\Lambda_b/\gamma_b]\}\\&\times\tfrac{1}{4}\pi^{1/2}\wp^2(Ku\Delta)^{-1}E_1E_2[\exp-(\omega-\nu_1)^2/(Ku)^2\\&+\exp-(\omega-\nu_2)^2/(Ku)^2]\cos(\pi z/L)\sin(\nu_2-\nu_1)t.\end{aligned}\tag{164}$$

From this expression one sees that the amplitude of the pulsations at frequency near Δ, relative to the dc change in population due to o.f. oscillation, should be

$$2E_1E_2(E_1^2+E_2^2)^{-1}(\gamma_a/\Delta)\cos(\pi z/L),$$

if, for simplicity, we neglect the Gaussian exponentials. Here again, it might be useful to use this phenomenon as a diagnostic tool while undertaking a systematic study of two-frequency operation.

16. CONNECTIONS WITH PREVIOUS CALCULATIONS

The basic paper in this field is, of course, that of Schawlow and Townes[1] who give expressions for threshold equivalent to (62). Townes[22] has also given an equation for linear pulling, as has Javan[23] for nonlinear pulling.

Oscillations of an optical maser involve the propagation of radiation in a nonlinear medium. Several papers have recently appeared which deal with this subject. For various reasons, these do not apply very closely to our particular problem. Thus, Bloembergen *et al.*[24] and Franken and Ward[25] treated harmonic generation which plays a relatively minor role for us. Teng and Statz[26] discussed low-order nonlinearities in a gaseous medium, but, as will be discussed below, their treatment of Doppler broadening is not adequate for our purposes. Also our model for radiation damping is more realistic than theirs which involves just one relaxation time τ, while our equations contain two decay constants γ_a and γ_b. To be sure, the combination $\gamma_{ab}=\frac{1}{2}(\gamma_a+\gamma_b)$ enters most equations, and this might be identified with τ^{-1}.

Among other publications which deal with maser theory are those of Wagner and Birnbaum[27] and of McCumber.[28] These papers consider to some extent the quantum nature of the electromagnetic field. They differ greatly in spirit and content from ours, and we will not attempt to make a comparison here. The work of Haken and Sauermann[29] and of Davis[30] seems much closer to ours, but there are significant differences in the models used, and in the appearance of our equations.

As mentioned in Sec. 4 an earlier calculation[8] applicable to an optical gaseous maser neglected complications arising from the atomic motions and multimode oscillation. It was then possible to work with a density matrix $\rho(\mathbf{r},t)$ characterizing an ensemble[7] of atoms at position $\mathbf{r}$ at time t which were excited at any time $t_0\leqslant t$. This obeyed

$$i\dot{\rho}=[\mathfrak{IC},\rho]-(i/2)(\Gamma\rho+\rho\Gamma)+i\Lambda,\tag{165}$$

which differs from (18) by the term containing a source matrix Λ describing the (slowly varying) rate densities of excitations Λ_a and Λ_b

$$\Lambda=\begin{pmatrix}\Lambda_a & 0\\ 0 & \Lambda_b\end{pmatrix}.\tag{166}$$

(In most applications Λ will be a diagonal matrix.)

It is possible to carry the calculation to higher order in the E_n for multimode oscillation without atomic motion by an iterative procedure in which we begin by neglecting any time dependence in the population difference $\rho_{aa}-\rho_{bb}$. In the rotating wave approximation, one of Eqs. (25) then gives a quasisteady-state solution for $\rho_{ab}(z,t)$

$$\begin{aligned}\rho_{ab}(z,t)=-\tfrac{1}{2}i(\wp/\hbar)\sum_\mu E_\mu U_\mu(z)\mathfrak{D}(\omega-\nu_\mu)\\\times(\rho_{aa}-\rho_{bb})\exp-i(\nu_\mu t+\varphi_\mu).\end{aligned}\tag{167}$$

Inserting this in another one of Eqs. (25) and again making a rotating wave approximation we find rate equations

$$\begin{aligned}\dot{\rho}_{aa}&=-\gamma_a\rho_{aa}+R(\rho_{bb}-\rho_{aa})+\Lambda_a,\\\dot{\rho}_{bb}&=-\gamma_b\rho_{bb}+R(\rho_{aa}-\rho_{bb})+\Lambda_b,\end{aligned}\tag{168}$$

where

$$\begin{aligned}R=\tfrac{1}{4}(\wp/\hbar)^2\sum_\lambda\sum_\mu E_\lambda E_\mu U_\lambda(z)U_\mu(z)\\\times[\mathfrak{D}(\omega-\nu_\mu)(\exp i(\nu_\lambda-\nu_\mu)t+i(\varphi_\lambda-\varphi_\mu))+\text{c.c.}].\end{aligned}\tag{169}$$

The "rate constant" R has pulsations for cases of multifrequency operation, and through Eq. (168) these would lead to pulsations in the populations ρ_{aa} and ρ_{bb} at all beat frequencies $\nu_\lambda-\nu_\mu$. If it were deemed necessary to continue the iteration procedure the pulsating population difference $\rho_{aa}-\rho_{bb}$ could be approximately evalu-

[22] C. H. Townes, *Advances in Quantum Electronics*, edited by J. Singer (Columbia University Press, New York, 1961), pp. 3–11.
[23] A. Javan, E. A. Ballik, and W. L. Bond, J. Opt. Soc. Am. **52**, 96 (1962).
[24] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. **127**, 1918 (1962).
[25] P. A. Franken and J. F. Ward, Rev. Mod. Phys. **35**, 23 (1963).
[26] C. L. Tang and H. Statz, Phys. Rev. **128**, 1013 (1962).
[27] W. G. Wagner and G. Birnbaum, J. Appl. Phys. **32**, 1185 (1961).
[28] D. E. McCumber, Phys. Rev. **130**, 675 (1963).
[29] H. Haken and H. Sauermann, Z. Physik **173**, 261 (1963); **176**, 47 (1963).
[30] L. W. Davis, Proc. Inst. Elec. Engrs. **51**, 76 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

ated from Eq. (168) and put in Eqs. (25) to obtain an improved Eq. (167).

If we neglect the pulsations in R, we obtain

$$R=\tfrac{1}{2}(\wp/\hbar)^2\gamma_{ab}\sum_\mu E_\mu^2[U_\mu(z)]^2\mathfrak{L}(\omega-\nu_\mu), \quad (170)$$

which is a plausible generalization of the rate of transitions previously obtained. In a steady state, the population difference implied by (168), (170) is

$$\rho_{aa}-\rho_{bb}=[(\Lambda_a/\gamma_a)-(\Lambda_b/\gamma_b)] \times[1+(2\gamma_{ab}R/(\gamma_a\gamma_b))]^{-1}, \quad (171)$$

which shows clearly the effect of o.f. saturation.

Since the rate constant (170) depends on position z through normal modes $U_n(z)$, the population difference (171) also depends on position, and may be said to have "holes" burned in it. Consequences of this for mode competition can be discussed along the lines followed in Sec. 10. Under most conditions the behavior will correspond to the weak-coupling case, although if the excitation density were such that $-N_2$ were larger than $\bar{N}$ the strong coupling case might be realized.

Combining Eqs. (165), (167), (171), and (41) we obtain a polarization

$$P(z,t)=\wp(\rho_{ab}+\rho_{ab}^*) \quad (172)$$

and the coefficients S_n and C_n which enter the amplitude and frequency determining Eqs. (13) and (14) are

$$S_n=-\frac{2}{L}(\wp^2/\hbar)\gamma_{ab}\mathfrak{L}(\omega-\nu_n)E_n \times\int_0^L dz[U_n(z)]^2N(z,t)[1+(2\gamma_{ab}R(z))/\gamma_a\gamma_b)]^{-1} \quad (173)$$

and

$$C_n=[(\omega-\nu_n)/\gamma_{ab}]S_n. \quad (174)$$

These expressions depend nonlinearly on the mode amplitudes E_n because of saturation. However, since we have already neglected the beat frequency pulsations of $\rho_{aa}-\rho_{bb}$ which lead to terms with Δ in the denominator, it might not be consistent to keep any terms in R which are off-resonance by more than about $|\omega-\nu_\mu|\sim(\gamma_{ab}\Delta)^{1/2}$.

We will now consider only single-frequency operation, for which our Eqs. (173)–(174) are essentially exact. Then a single summand $\mu=n$ contributes to R and we find

$$S_n=-\frac{2}{L}\wp^2\hbar^{-1}\gamma_{ab}E_n\int_0^L dz[U_n(z)]^2N(z)[\gamma_{ab}^2+(\omega-\nu_n)^2 +\gamma_{ab}^2(\wp E_n)^2[U_n(z)]^2/(\hbar^2\gamma_a\gamma_b)]^{-1}. \quad (175)$$

The integral may be easily calculated if $N(z)$ is a slowly varying function of position, and we find

$$S_n=-\wp^2\hbar^{-1}\gamma_{ab}\bar{N}\mathfrak{L}(\omega-\nu_n)f(w)E_n, \quad (176)$$

where

$$w=[(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)]\gamma_{ab}^2\mathfrak{L}(\omega-\nu_n) \quad (177)$$

and

$$f(w)=(2/w)[1-(1+w)^{-1/2}]. \quad (178)$$

If Eq. (175) for S_n is expanded to third order in E_n, we obtain a result in agreement with that given by Eqs. (56), (72) for the single-frequency case with no atomic motion. It should be noted that whether (175) is expanded or not, the amplitude of oscillation has a maximum for resonant tuning and falls off monotonically with detuning. There is no indication here that the double maximum of Sec. 8 might be a spurious one which arises from the neglect of fifth or higher order terms.

17. DISCUSSION OF DOPPLER BROADENING

The effect of atomic motion upon our equations is rather curious and warrants discussion which, for simplicity, will be given for single-frequency operation. As we have seen in Sec. 16, the optical properties of the medium may be described by a nonlinear susceptibility

$$\chi(\omega-\nu_n, E_n)=P_n/(\epsilon_0E_n). \quad (179)$$

It would perhaps be plausible to hope that the following simple recipe would take atomic motion into account. Because an atom moving with velocity v sees a Doppler shifted frequency, in the laboratory frame of reference it effectively has its resonance frequency shifted by $\omega v/c$. The effective susceptibility ought then to be

$$\chi_{\text{eff}}=\int_{-\infty}^{\infty}dvW(v)\chi(\omega-\omega(v/c)-\nu_n, E_n), \quad (180)$$

which can easily be expressed in terms of the $Z(\zeta)$ function which is discussed in Sec. 6. The effective damping constant becomes

$$\gamma_{ab}[1+(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)]^{1/2}$$

instead of γ_{ab}, and when this is much smaller than the Doppler parameter Ku, the line shape should be Gaussian with a normal Doppler width.

The above prescription is incorrect except to the first order in E_n if standing waves rather than running waves are involved in any way. This follows from a study of Eq. (70) contributing to the third-order polarization which involves a threefold integration over times τ', τ'', and τ'''. For single-frequency oscillation, the integrand contains

$$U_n(z-v\tau')U_n(z-v\tau'-v\tau'')U_n(z-v\tau'-v\tau''-v\tau''') \times\exp[-i(\omega-\nu_n)\tau']\exp[\pm i(\omega-\nu_n)\tau'''].$$

Each of the U_n's is a standing wave which can be expressed as the sum of two running waves like $\exp\pm i(Kz-v\tau')$, etc. The physical interpretation is that in order to contribute to $P_n^{(3)}$ at time t an atom

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

has to interact three times with the o.f. field: first at $t'''=t-\tau'-\tau''-\tau'''$, then at $t''=t-\tau'-\tau''$, and finally at $t'=t-\tau'$. At each of these interactions the atom has a "choice" of interacting with one or the other of the two running waves. The τ dependence of a typical term of the integrand of (71) is then

$$\exp-\gamma_{ab}(\tau'+\tau''')-\gamma_a\tau''\exp i(\nu_n-\omega)(\tau'\pm\tau''')\pm iK\nu\tau' \pm iK\nu(\tau'+\tau'')\pm iK\nu(\tau'+\tau''+\tau''').$$

The physical consequence of the appearance of terms involving $\pm K\nu$ in the exponent is that each interaction involves a Doppler shift. Since we are interested in the case of a large Doppler width Ku, such a shift can take a given interaction very far from resonance even if $\omega=\nu_n$. This is expressed by the destructive interference of contributions to the integral at various value of τ', τ'', and τ'''. The interference will be least when the accumulated Doppler phase angle

$$\pm K\nu\tau'\pm K\nu(\tau'+\tau'')\pm K\nu(\tau'+\tau''+\tau''')$$

is zero. In order to obtain a nonvanishing spatial Fourier projection (8) for $P_n{}^{(3)}$, the choices $\pm$ can not be all alike. The six remaining possibilities for the phase angle are

$$\pm K\nu[\tau'+(\tau'+\tau'')-(\tau'+\tau''+\tau''')]=\pm K\nu(\tau'-\tau'''),$$
$$\pm K\nu[\tau'-(\tau'+\tau'')+(\tau'+\tau''+\tau''')]=\pm K\nu(\tau'+\tau'''),$$

and

$$\pm K\nu[-\tau'+(\tau'+\tau'')+(\tau'+\tau''+\tau''')]= \pm K\nu(\tau'+2\tau''+\tau''').$$

As seen in Sec. 7, only the first two possibilities are able to lead to vanishing interference, and then only when $\tau'=\tau'''$.

Physically, one may say that a dominant type of process involves three interactions: first, one with a right (left) running wave at t''', then one with a left (right) running wave at t'', and finally one with a left (right) running wave at t', with the time intervals obeying $t-t'=t''-t'''$ so that the accumulated Doppler phase angle

$$\pm[K\nu(t-t')+K\nu(t-t'')-K\nu(t-t''')]$$

cancels out at time t.

The above cancellation of Doppler interference would not occur if waves running in only one direction were present. The nonlinear terms are much less broadened and weakened by Doppler effect for a standing wave maser oscillator than would be inferred from a study of nonlinear propagation alone. The double peak in the power as a function of tuning met in Sec. 8 can occur only because β_n (saturation) is not as much Doppler broadened as α_n (linear gain profile).

18. HOLE BURNING

In his discussion of maser action Bennett[31] has made use of the notion of "hole burning." Since it aids the physical understanding of the rather complicated equations, we will now show how this phenomenon is described in the present work. As Bennett's treatment does not bring in the population pulsations of Sec. 15, we will base our discussion on the simplified theory of Sec. 16 in which pulsations of population were neglected.

In Sec. 16 the atoms had zero velocity. It is possible to generalize the discussion for the case of atoms having a velocity distribution $W(v)$ at the cost of further fairly plausible assumptions which are no worse than approximations already made. We deal first with those atoms which have a definite velocity v and which were excited at z_0 at time t_0. The perturbation experienced by such an atom is

$$V(t)=-(\wp/\hbar)\sum_\mu E_\mu U_\mu(z_0+v(t-t_0))\cos(\nu_\mu t+\varphi_\mu) \quad (181)$$

so that instead of seeing fields at frequencies ν_μ the moving atom sees fields at twice as many frequencies, i.e., $\nu_\mu\pm Kv$. The rate concept approach of Sec. 16 can not be used since the atoms characterized by different values of z_0 and t_0 experience different perturbations, i.e., the *phases* are not the same for the various members of the ensemble of atoms arriving at z at time t. It is plausible, however, to estimate the effect of saturation on the population difference $\rho_{aa}(v,t)-\rho_{bb}(v,t)$ by using an equation like (171) with a perturbation $V(t)$ given by (181) but with the terms involving z_0 and t_0 omitted. The replacement for the velocity-dependent rate constant $R(v)$ is then plausibly

$$R(v)=\tfrac{1}{8}\gamma_{ab}(\wp/\hbar)^2\sum_\mu E_\mu{}^2[\mathcal{L}(\nu_\mu-\omega+Kv) +\mathcal{L}(\nu_\mu-\omega-Kv)]. \quad (182)$$

If $v=0$, this reduces to the space average of Eq. (170). The corresponding population difference for atoms having velocity v is then

$$\rho_{aa}(v,t)-\rho_{bb}(v,t) =W(v)[(\Lambda_a/\gamma_a)-(\Lambda_b/\gamma_b)] \times[1+2\gamma_{ab}R(v)/(\gamma_a\gamma_b)]^{-1}. \quad (183)$$

A plot of (183) against v would show the assumed velocity distribution with "holes" burned into it due to o.f. saturation effects. These holes would be appreciable whenever $R(v)$ became comparable to $\gamma_a\gamma_b/(\gamma_a+\gamma_b)$ which could occur near $Kv=\pm(\nu_\mu-\omega)$, so that *two* holes could be burned for each cavity mode in oscillation. At a resonance, where $\nu_n=\omega$, the two holes for the nth mode would merge and reinforce one another. The holes in the velocity distribution would have been

[31] W. R. Bennett, Jr., Appl. Opt. Suppl. 1, 24–61 (1962), especially pp. 58–59. It should be noted that the holes are in first instance burned in the curve of population difference versus velocity, and only indirectly in a curve of gain versus frequency.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

seen in Sec. 15 if the integration over velocities had not been carried out so soon.

With the above approximate expression for $\rho_{aa}(v)-\rho_{bb}(v)$ we may use Eqs. (25) to calculate a pseudo-first-order value of $\rho_{ab}(z,v,t)$ and $P(z,v,t)$ in the manner of Sec. 16. The result will be the same as before, but with a velocity dependent reduction factor

$$[1+2\gamma_{ab}R(v)/(\gamma_a\gamma_b)]^{-1}$$

to express the effect of saturation, where $R(v)$ is given by (182). There are similar reductions in the related functions $S_n(v)$ and $C_n(v)$. The coefficients S_n and C_n which enter the equations of self-consistency [(13),(14)] result from integration of these quantities over velocity. Thus

$$S_n=-\wp^2\hbar^{-1}\bar{N}\gamma_{ab}E_n\int_{-\infty}^{\infty}dvW(v)\mathfrak{L}(\omega-\nu_n+Kv)\times[1+2\gamma_{ab}R(v)/(\gamma_a\gamma_b)]^{-1}. \quad (184)$$

For single-mode operation, the rate $R(v)$ in the denominator of (184) contains the two terms given by Eq. (182). If an expansion is made to first order in $R(v)$ the integration over v can be done easily for a Maxwellian velocity distribution in the limit of large Doppler width Ku, and one gets results equivalent to those of Sec. 8. The possible dip in output power as a function of cavity tuning in single-mode operation can be interpreted as a consequence of the merging of the two holes at $Kv=\pm(\nu_n-\omega)\to 0$.

For multimode operation a similar calculation can be made with the complete expression (183) for $R(v)$. This will give the dominant terms of the equations of Sec. 9 for a Maxwellian velocity distribution and large Doppler width. However, the frequency locking terms involving ψ will be missing.

For two-mode oscillation, (184) leads to especially strong mode competition attributable to hole burning when $Kv=\nu_2-\omega=\omega-\nu_1$, i.e., the traveling wave along $+z$ for mode 2 and the traveling wave along $-z$ for mode 1 are both in approximate resonance with the atomic transition frequency for an atom having velocity v. This effect can be correlated with the peak in θ for $\omega=\nu_{12}$ mentioned in Sec. 10.

19. APPROXIMATE HIGHER ORDER THEORY FOR SINGLE MODE OPERATION

It would be possible, but quite tedious, to extend the calculations of the text to fifth and higher order for the single-frequency case. The simpler approximate procedure outlined below may serve in the absence of such calculations. It was mentioned above that an expansion of (184) to first order in $R(v)$ reproduces the equations of single-mode operation correct to third order in E_n. If this expansion is not made one may hope to have equations which are valid for stronger signals. The v integration is complicated, and we will content ourselves here with two special cases (a) $Ku\gg|\nu_n-\omega|\gg\gamma_{ab}$ and (b) $\nu_n=\omega$, $Ku\gg\gamma_{ab}$. In the former case, one finds approximately

$$S_n=-\pi^{1/2}\wp^2\bar{N}E_n/(\hbar Ku)[1+\tfrac{1}{4}(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)]^{1/2}, \quad (185)$$

while at resonance, when $\nu_n=\omega$,

$$S_n=-\pi^{1/2}\wp^2\bar{N}E_n/(\hbar Ku)[1+\tfrac{1}{2}(\wp E_n)^2/(\hbar^2\gamma_a\gamma_b)]^{1/2}, \quad (186)$$

and the merging of the two holes shows up in a simple manner through the doubling of the term expressing the effects of saturation. It will be remembered that a similar doubling of the coefficient β_n took place in Sec. 8 and was responsible for the possible dip in maser output versus cavity tuning. Although the more general behavior of output versus tuning implied by Eqs. (185)–(186) should be qualitatively correct, it must be remembered that rather uncontrolled approximations have been made in their derivation.

20. EXCITATION OF LOWER MASER LEVEL BY SPONTANEOUS DECAY OF UPPER MASER STATE

It was mentioned in Sec. 6 that the lower maser level could, at least in part, be excited by spontaneous emission from the upper. For the present this complication will be treated only in an approximation in which the rate concept is valid. For simplicity we ignore atomic motion, although the work of Secs. 16 and 19 suggests how this could be allowed for approximately. We write rate equations like (168)

$$\begin{aligned}\dot{\rho}_{aa}&=-\gamma_a\rho_{aa}+R(\rho_{bb}-\rho_{aa})+\Lambda_a,\\ \dot{\rho}_{bb}&=-\gamma_b\rho_{bb}+R(\rho_{aa}-\rho_{bb})+\Lambda_b+f\gamma_a\rho_{aa},\end{aligned} \quad (187)$$

where the extra source term $f\gamma\rho_{aa}$ describes the effects of radiative cascade excitation of b from a assuming that a fraction f of the decays from a are to b. The Λ's describe the uncorrelated excitation of the two levels. In a steady state one finds a population density difference

$$\rho_{aa}-\rho_{bb}=[(\Lambda_a/\gamma_a)(1-f(\gamma_a/\gamma_b))-(\Lambda_b/\gamma_b)]\times[1+R\{\gamma_a(1-f)+\gamma_b\}/(\gamma_a\gamma_b)]^{-1}, \quad (188)$$

which should be compared to Eq. (171). It will be seen that the effect of a nonvanishing branching ratio f is merely to change the unsaturated population difference (obtained for $R=0$), and also to modify the value of R for which a given degree of saturation would be obtained. The saturation parameter of Sec. 8 will be modified in an obvious fashion. Thus if $f=0$, a value of rate $R=\frac{1}{2}\gamma_a\gamma_b/\gamma_{ab}$ would cause 50% saturation, while if $f=1$ the value would be $R=\gamma_a$. It should be recalled that the dominant part of the third-order terms $S_n^{(3)}$ and $C_n^{(3)}$ are direct manifestations of saturation phenomena. At the present state of maser art, the decay constants γ_a and γ_b are not well enough known

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

for the effect of a nonvanishing value of f to be easily seen.

When population pulsations are taking place there will be a correlated time-dependent excitation of the lower level by cascade. It is possible that more interesting consequences than those obtained would result, and it is hoped to explore this possibility in a later paper.

21. OTHER SOURCES OF BROADENING

For some kinds of line broadening, especially in certain solid-state optical masers, one could adopt the recipe proposed in Sec. 17, and rejected for the case of Doppler broadening. If the effect of environment could be described by a distribution function for the atomic resonance frequencies ω, an averaged nonlinear susceptibility could be used. This could also be done for the case of isotopic mixtures of the active atoms in gaseous masers.

Although γ_a and γ_b were introduced into our equations to describe spontaneous radiative decay of the states a and b, it is plausible that such phenomenological decay constants might also describe certain kinds of collision broadening. In that case, the γ's would be functions of the pressure.[32] A more detailed discussion of collision broadening for a gaseous optical maser will be given in another paper.

[32] Evidence for such a dependence has recently been obtained by Javan and Szöke, Ref. 16.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Correlation Effects in Many Fermion Systems. II. Linked Clusters*

HUGH P. KELLY
Department of Physics, University of California at San Diego, La Jolla, California
(Received 31 January 1964)

In a previous paper a set of coupled equations was derived for the ground-state wave function and energy of a finite system of interacting Fermions. The equations are now modified so as to be more applicable to systems in which the number of particles becomes large. The resulting equations are shown to be equivalent to those obtained from many-body perturbation theory.

I. LINKED CLUSTERS

IN a previous paper,[1] a set of coupled equations was derived for the ground-state wave function and energy of a finite system of interacting Fermions. The wave function was expanded in terms of multiple-particle excitations on an uncorrelated zero-order state. The total energy E of the system appeared in the resulting equations and it was pointed out that this restricts the application of these equations to finite systems; in general, the restriction is to systems of small N. In the equations, the amplitudes for one-particle excitations are coupled to those for two-particle and three-particle excitations. The two-particle amplitudes are coupled to those for one-particle, three-particle, and four-particle excitations, and similarly for higher particle excitations. It was mentioned in I that it might be reasonable to approximate four-particle excitation terms, for example, as products of independent two-particle excitations.

It is shown here that four-particle terms involving two independently propagating pairs enter the equations in such a way as to eliminate the dependence of the two-particle excitation equations on the total energy E, and similarly for the other excitations.[2] Explicit inclusion of products of independent excitations yields the equations of the linked cluster expansion. The resulting equations are shown to be the same as those obtained from many-body perturbation theory as formulated by Brueckner[3] and by Goldstone.[4]

In I, the ground-state wave function is expanded as

$$|\psi\rangle = |\Phi_0\rangle + \sum_{\alpha,k} f(k;\alpha)\eta_k^+\eta_\alpha|\Phi_0\rangle + \sum_{\alpha,\beta,k,k'} f(kk';\alpha\beta)\eta_k^+\eta_{k'}^+\eta_\beta\eta_\alpha|\Phi_0\rangle + \cdots . \quad (1)$$

The unperturbed solution $|\Phi_0\rangle$ is a determinant composed of the N single-particle states which are lowest in energy.

Equations are then derived for $f(k;\alpha)$ and $f(kk';\alpha\beta)$ by inserting $|\psi\rangle$ from Eq. (1) into

$$H|\psi\rangle = E|\psi\rangle , \quad (2)$$

where H is written in the usual second-quantized form.[1]

* Work supported in part by the U. S. Atomic Energy Commission.

[1] H. P. Kelly and A. M. Sessler, Phys. Rev. 132, 2091 (1963), hereafter referred to as I.

[2] I am indebted to Dr. A. M. Sessler for stressing the desirability of including products of independent pair excitations in the four-particle excitations, so as to make the resulting equations more applicable to systems of large N.

[3] K. A. Brueckner, in *The Many Body Problem*, edited by C. DeWitt (John Wiley & Sons, Inc., New York, 1958).

[4] J. Goldstone, Proc. Roy. Soc. (London) A239, 267 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

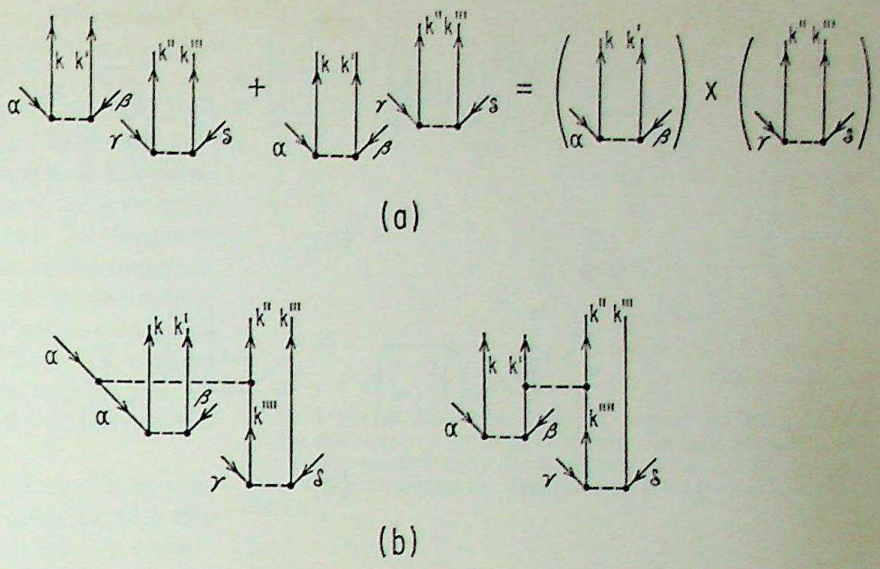


FIG. 1. (a) In perturbation theory two independent pair excitations factor when both time orderings are considered. (b) Typical terms which always link any two excited pairs.

An approximate solution for $f(kk';\alpha\beta)$ is given in Eq. (15) of I, in this paper referred to as Eq. (I.15),

$$f(kk';\alpha\beta)=(\epsilon_\alpha+\epsilon_\beta-\epsilon_k-\epsilon_{k'}-\langle(\alpha\beta)_{ex}|v|\alpha\beta\rangle +E-E_{HF})\langle(kk')_{ex}|v|\alpha\beta\rangle. \quad (3)$$

If we neglect the term $E-E_{HF}$, then this expression is that which is obtained from Rayleigh-Schrödinger perturbation theory when we use a Hartree-Fock basis and include the first-order term and all higher order diagonal hole-hole interaction terms.[5,6] However, for large N the term $E-E_{HF}$ may become so large as to invalidate the equations of I. Such difficulties have been described by Brueckner in his comparison of Rayleigh-Schrödinger and Brillouin-Wigner perturbation theories.[3]

We now show that the term $E-E_{HF}$ may be removed by considering the coupling of four-particle excitations in Eq. (I.11) for $f(kk';\alpha\beta)$. In the equation for $f(kk';\alpha\beta)$, coupling with four-particle excitations adds the term

$$\sum_{\gamma\delta k''k'''}\langle(\gamma\delta)_{ex}|v|k''k'''\rangle f(kk'k''k''';\alpha\beta\gamma\delta) \quad (4)$$

to the left-hand side of Eq. (2) or Eq. (I.11). The coefficients $f(kk'k''k''';\alpha\beta\gamma\delta)$ are composed of parts in which two pairs propagate independently and of remaining parts involving more complicated interactions among the four particles.

The approximation of regarding $f(kk'k''k''';\alpha\beta\gamma\delta)$ solely as a product of two pair excitations has been discussed by Brenig and Sinanoğlu.[7] A justification for this decomposition is found in perturbation theory where the lowest order four-particle excitation is given by the product of two pair excitations when we consider both possible time orderings as in Fig. 1(a). Those parts of $f(kk'k''k''';\alpha\beta\gamma\delta)$ in which two pairs of particles propagate independently should be written as products of $f(kk';\alpha\beta)$. It may be noted that there are always linked terms connecting any two excited pairs, however, as shown in Fig. 1(b).

When product pairs such as $f(kk'';\alpha\gamma)f(k'k''';\beta\delta)$ are inserted into Eq. (4) and used in the equation for $f(kk';\alpha\beta)$, than linked terms result as shown in Fig. 2.

When we consider the products

$$f(kk';\alpha\beta)\sum f(k''k''';\gamma\delta),\quad \gamma\neq\alpha,\beta\quad k''\neq k,k'\quad \delta\neq\alpha,\beta\quad k'''\neq k,k', \quad (5)$$

these give us terms

$$f(kk';\alpha\beta)\sum\langle\gamma\delta|v|(k''k''')_{ex}\rangle f(k''k''';\gamma\delta),\quad \gamma\neq\alpha,\beta\quad k''\neq k,k'\quad \delta\neq\alpha,\beta\quad k'''\neq k,k' \quad (6)$$

on the left-hand side of Eq. (2).

In Eq. (I.8) it is shown that

$$E-E_{HF}=\sum_{\alpha,\beta,k,k'}\langle\alpha\beta|v|(kk')_{ex}\rangle f(kk';\alpha\beta). \quad (7)$$

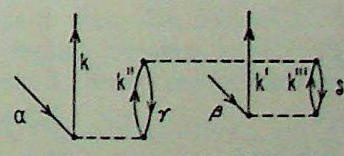


FIG. 2. Diagram illustrating how H may couple two independent pair excitations with a single pair excitation. There is also a diagram in which the time ordering of the first two interactions is reversed. This term enters the equation for $f(kk';\alpha\beta)$.

[5] H. P. Kelly, Phys. Rev. 131 684 (1963).

[6] Hartree-Fock single-particle states are assumed, although they are not essential for this discussion.

[7] W. Brenig, Nucl. Phys. 4, 363 (1957); O. Sinanoğlu, J. Chem. Phys. 36, 706 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

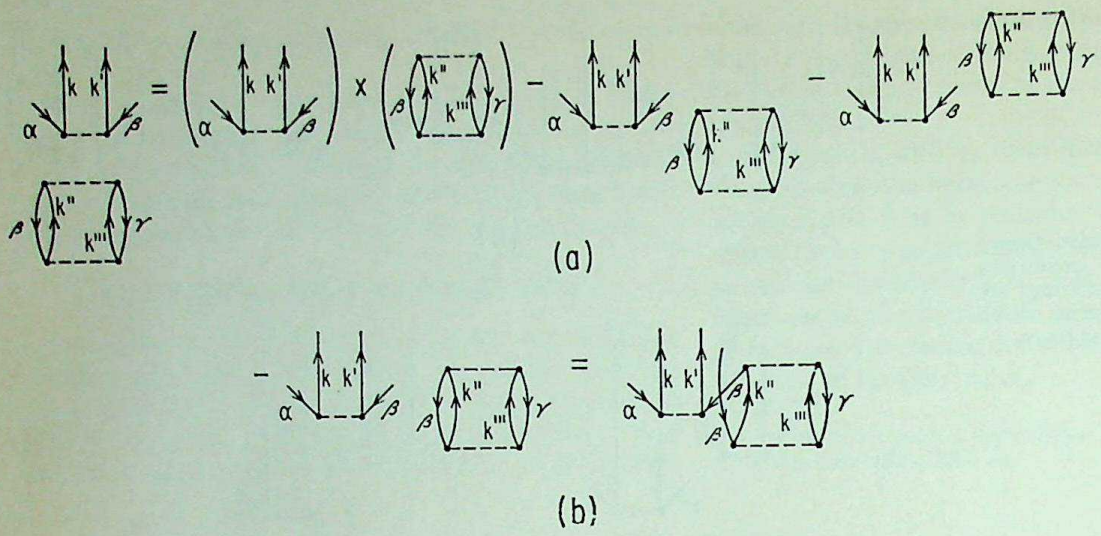


FIG. 3. (a) Diagram factorization when the two disconnected parts have one or more particle or hole lines in common. The two subtracted diagrams are really linked diagrams as shown in (b).

Equation (6) is then rewritten

$$f(kk';\alpha\beta)[E-E_{HF}-\sum_{\delta,k'',k'''}\langle\alpha\delta|v|(k''k''')_{ex}\rangle(k''k''';\alpha\delta)$$
$$-\sum_{\gamma\neq\alpha,k'',k'''}\langle\gamma\beta|v|(k''k''')_{ex}\rangle f(k''k''';\gamma\beta)$$
$$-\sum_{k'''';\ \gamma\neq\alpha,\beta;\ \delta\neq\alpha,\beta}\langle\gamma\delta|v|(kk''')_{ex}\rangle f(kk''';\gamma\delta)$$
$$-\sum_{k''\neq k;\ \gamma\neq\alpha,\beta;\ \delta\neq\alpha,\beta}\langle\gamma\delta|v|(k''k')_{ex}\rangle f(k''k';\gamma\delta)]. \quad (8)$$

When the term of Eq. (8) is added to the left-hand side of Eq. (I.11), $f(kk';\alpha\beta)(E-E_{HF})$ cancels on both sides of Eq. (I.11).

At this point we now have the linked cluster expansion for two-particle excitations.[4] Equation (I.11) is also now nonlinear but in many cases this should not cause any difficulty.[5] The remaining terms are not proportional to N like $E-E_{HF}$; they will be given a simple physical interpretation. The denominators in Eqs. (I.15) and (I.16) are now modified by replacing the terms $E-E_{HF}$ by the summations in Eq. (8) with positive signs. For example, Eq. (I.15), an approximate equation for $f(kk';\alpha\beta)$, becomes

$$f(kk';\alpha\beta)$$
$$=[\epsilon_\alpha+\epsilon_\beta-\epsilon_k-\epsilon_{k'}-\langle(\alpha\beta)_{ex}|v|\alpha\beta\rangle$$
$$+\sum_{\delta,k'',k'''}\langle\alpha\delta|v|(k''k''')_{ex}\rangle f(k''k''';\alpha\delta)$$
$$+\sum_{\gamma\neq\alpha,k'',k'''}\langle\gamma\beta|v|(k''k''')_{ex}\rangle f(k''k''';\gamma\beta)$$
$$+\sum_{k'''';\ \gamma\neq\alpha,\beta;\ \delta\neq\alpha,\beta}\langle\gamma\delta|v|(kk''')_{ex}\rangle f(kk''';\gamma\delta)$$
$$+\sum_{k''\neq k;\ \gamma\neq\alpha,\beta;\ \delta\neq\alpha,\beta}\langle\gamma\delta|v|(k''k')_{ex}\rangle f(k''k';\gamma\delta)]^{-1}$$
$$\times\langle(kk')_{ex}|v|\alpha\beta\rangle. \quad (9)$$

The new terms appearing in the denominator of Eq. (9) which refer to states $|\alpha\rangle$ and $|\beta\rangle$ combine with the single-particle energies ϵ_α and ϵ_β and with the diagonal hole-hole interaction term $\langle(\alpha\beta)_{ex}|v|\alpha\beta\rangle)$ to give an effective two-body energy for $|\alpha\beta\rangle$. When the denominator also includes the particle-particle and exclusion principle violating (EPV) hole-particle terms of Eq. (I.17), then the denominator becomes approximately the difference between the two-body energies of the states $|\alpha\beta\rangle$ and $|kk'\rangle$. However, the two-body energy of $|kk'\rangle$ is not given so accurately as that of $|\alpha\beta\rangle$ because the interaction of the excited particles with unexcited particles is accounted for only in the Hartree-Fock approximation. However, pair-correlation energy terms of a ground-state particle with the other ground-state particles are now included. The last two terms in the denominator of Eq. (9) which refer to states $|\gamma\delta\rangle$ account for the fact that correlation terms for ground-state particles which involve excited states $|k\rangle$ or $|k'\rangle$ will be eliminated by the Pauli principle when there are other particles excited into $|k\rangle$ or $|k'\rangle$.

As a simple example of this result we consider the beryllium atom and let α, β, γ, δ refer to the $2s^+$, $2s^-$, $1s^+$, and $1s^-$ single-particle states, respectively. In Ref. 5 it was shown that the $1s-2s$ interactions are small compared to $1s-1s$ and $2s-2s$ interactions and so $f(kk';1s2s)$ is omitted in the denominators. Also the excited single-particle states used in Ref. 5 were all in the continuum and so the sums in Eq. (9) which do not run over both continuum states vanish because, as shown in Ref. 5, $\Sigma=(R_0/\pi)\int dk$ and continuum states have normalization $(2/R_0)^{1/2}$. Equation (9) becomes

$$f(kk';\alpha\beta)$$
$$=(\epsilon_\alpha+\epsilon_\beta-\langle(\alpha\beta)_{ex}|v|\alpha\beta\rangle+\sum_{k'',k'''}\langle\alpha\beta|v|(k''k''')_{ex}\rangle$$
$$\times f(k''k''';\alpha\beta)-\epsilon_k-\epsilon_{k'})^{-1}\times\langle(kk')_{ex}|v|\alpha\beta\rangle. \quad (10)$$

In the denominator of Eq. (10) we have the effective two-particle energy for the state $|\alpha\beta\rangle$. The term

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$-\langle(\alpha\beta)_{ex}|v|\alpha\beta\rangle$ corrects for the inclusion of the Coulomb interaction of $|\alpha\rangle$ and $|\beta\rangle$ in both single-particle energies ϵ_α and ϵ_β. The term $\sum_{k''k'''}\langle\alpha\beta|v|(k''k''')_{ex}\rangle f(k''k''';\alpha\beta)$ accounts for second and higher order interactions among the particles in states $|\alpha\rangle$ and $|\beta\rangle$. In the calculations of Ref. 5 it was found that omission of this term would have resulted in approximately a 10% error in the correlation energy among the two $2s$ electrons When there are bound "excited" states in the complete set of single-particle states then the energy denominators will also include terms from the last two summations in the denominator of Eq. (9). These correlation terms which we have been considering may be related to a type of "rearrangement diagram" considered by Brueckner and Goldman.[8]

The equation for one-particle excitations $f(k;\alpha)$ also has the term $E-E_{HF}$ removed when coupling with the three-particle excitation term is considered. We write

$$f(kk'k'';\alpha\beta\gamma)=f(k;\alpha)\sum_{k',k''\neq k,\beta,\gamma\neq\alpha}f(k'k'';\beta\gamma)+\cdots. \quad (11)$$

The result is that we replace $(E-E_{HF})$ in Eq. (I.10) by

$$\sum_{\gamma,k',k''}\langle\alpha\gamma|v|(k'k'')_{ex}\rangle f(k'k'';\alpha\gamma)+\sum_{\beta,\gamma\neq\alpha,k''}\langle\beta\gamma|v|(kk'')_{ex}\rangle f(kk'';\beta\gamma). \quad (12)$$

II. CONNECTION WITH PERTURBATION THEORY

The results which we have just obtained may also be derived by a consideration of the perturbation expansion.[3,4] We begin by examining the factorization of unlinked diagrams as shown in Fig. 3(a). The two apparently disconnected parts of each diagram are assumed to have one or more hole or particle lines in common and the two diagrams on the right of Fig. 3(a) must not be considered as "unlinked." They are really "linked" as is shown in Fig. 3(b). The first term on the right-hand side of Fig. 3(a) is eliminated by the usual cancellation of unlinked clusters.[3,4] Similar factorizations were considered in Ref. 5 where the two diagrams on the right of Fig. 3(a) were called third-class EPV diagrams. These diagrams were summed in an approximate way, although with sufficient accuracy for the numerical calculations of that paper. However, these two diagrams on the right of Fig. 3(a) may be summed exactly by noticing that both possible time orderings appear[9] or by merely writing down the algebraic expressions and adding.[8] For example, in the first subtracted

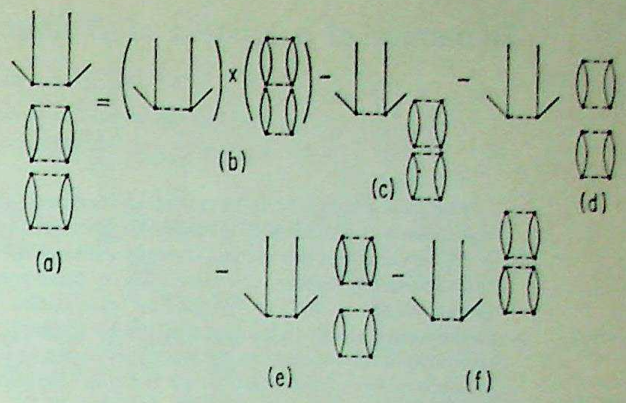


FIG. 4. Factorization of an unlinked diagram. It is assumed that all the disconnected parts have at least one hole or particle line in common so that there are correction terms (c), (d), (e), and (f) to the factorization (b). Diagrams (c) and (d) must be factored further.

diagram of Fig. 3(a), shown in Fig. 3(b), the product of the energy denominators is $(B(A+B)A)^{-1}$, where B is the denominator of the closed part considered separately and A is the denominator of the other part. The second subtracted diagram of Fig. 3(a) has the same matrix elements and a contribution $(A(A+B)A)^{-1}$ from the denominators. The sum of these expressions is $(A^2B)^{-1}$, so the two subtracted expressions add to give a result which is the first correction term obtained in an expansion of the linked part with a shifted denominator. In other words, it is the second term obtained in an expansion of the denominator of the following expression:

$$(\epsilon_\alpha+\epsilon_b-\epsilon_k-\epsilon_{k'}+\langle bc|v|(k''k''')_{ex}\rangle(\epsilon_b+\epsilon_c-\epsilon_{k''}-\epsilon_{k'''})^{-1}\times\langle k''k'''|v|bc\rangle)^{-1}\langle(kk')_{ex}|v|ab\rangle. \quad (13)$$

The higher order terms are obtained from a consideration of the factorizations of diagrams of Fig. 4. Such factorizations were shown in more detail in Ref. 5. Correction terms resulting from the factorization of the diagrams in Figs. 4(c) and 4(d) and from the diagrams of 4(e) and 4(f) add to give the second correction term in an expansion of the denominator of Eq. (13) and the first correction term which would result from using the appropriately shifted denominator in the correction term of the denominator of Eq. (13). The result of such factorizations then gives the summation terms in the denominator of Eq. (9) or, in a more simple case, it produces the shift in energy denominators as shown in Eq. (10).

ACKNOWLEDGMENTS

I would like to thank Dr. Andrew M. Sessler for his encouragement and for very helpful suggestions and Professor Keith A. Brueckner for his support and for a helpful discussion.

[8] K. A. Brueckner and D. T. Goldman, Phys. Rev. 117, 207 (1960).
[9] H. A. Bethe, B. H. Brandow, and A. G. Petschek, Phys. Rev. 129, 225 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Dynamics of a System of N Atoms Interacting With a Radiation Field

E. R. Buley and F. W. Cummings*
Aeronutronic Division, Philco Corporation, Newport Beach, California
(Received 4 December 1963)

The time-dependent behavior of an idealized laser is analyzed. The analysis avoids the use of the rate equations and includes N-particle "superradiant" effects. Perturbation theory is not used; the dynamics of N two-level atoms interacting strongly with a single-radiation field mode yields four coupled nonlinear equations which are integrated numerically for some special cases appropriate to lasers operating under ordinary power output as well as "Q-spoiling" conditions. Loss effects are grafted onto the dynamical equations phenomenologically. It is seen that under appropriate conditions, the laser is expected to emit one or more appreciable subsequent output bursts following an initial giant pulse, and that the subsequent emissions may be emitted by a population which is below inversion.

I. INTRODUCTION

THE continuing refinement of experimental techniques as well as of crystal growing techniques allows the mechanism underlying the production of coherent bursts of radiation, such as produced in lasers, to be probed in more detail. In the past, the analysis of the time dependence of a system of atoms coupled to a radiation field has most frequently been based on a rate equation, or energy balance approach.[1,2] While this approach is satisfactory in many situations, the rate equations ignore phase effects between the dipole moment of the radiators and the radiation field. As will be pointed out below, a more complete analysis can lead to interesting effects in lasers under certain conditions; in particular, to the extraction of coherent radiation from a population which is below inversion.

The model is essentially semiclassical, since the expectation values of products of field and atomic operators are factored into a product of field and atom expectation values. The results of this factorization do not differ significantly from the electrodynamic treatment for these purposes, and it is of some interest to note that "superradiant" spontaneous emission,[3] including the correct Einstein B coefficient, follows naturally from the resulting equations.

Phonomenological terms are added to the basic set of equations, in a way familiar from the field of magnetic resonance, to describe the loss of radiation from the resonant structure, the "dephasing" of the total dipole moment and the effect of pumping radiation. The equations are cast into a form amenable to numerical solution and a three-level, ruby prototype laser is discussed under various conditions such as Q spoiling[4] and variation of the material parameters.

II. MATHEMATICAL FORMULATION

Consider an N-atom system in which only two levels of each atom interact appreciably with the resonant modes of the electromagnetic (e.m.) field. In the dipole approximation, the Hamiltonian for the system, with zero-point energy of the e.m. field subtracted, is taken as

$$\mathcal{H}=\sum_i \hbar\Omega^i R_3{}^i+\sum_\lambda \hbar\omega_\lambda a_\lambda{}^+ a_\lambda + \sum_{\lambda,i}\hbar K_\lambda{}^i(a_\lambda{}^+ R_-{}^i+a_\lambda R_+{}^i)\,. \quad \text{(II.1)}$$

The operators $R_3{}^i$, $R_\pm{}^i$, following Dicke's notation,[3] in the space of the jth atom satisfy the commutation rules

$$[R_3{}^i, R_\pm{}^{i'}]=\pm R_\pm{}^i\delta_{ii'}\,, \quad \text{(II.2a)}$$

$$[R_\pm{}^i, R_\mp{}^{i'}]=2R_3{}^i\delta_{ii'}\,, \quad \text{(II.2b)}$$

respectively, and the field creation and destruction operators $a_\lambda{}^+$ and a_λ of the λth mode satisfy the commutation rules

$$[a_\lambda, a_{\lambda'}{}^+]=\delta_{\lambda\lambda'}\,, \quad \text{(II.3a)}$$

$$[a_\lambda, a_{\lambda'}]=0\,. \quad \text{(II.3b)}$$

The electric and magnetic fields are assumed to have the normal mode expansions

$$\mathbf{E}(\mathbf{x},t)=-\sum_\lambda(2\pi\hbar\omega_\lambda)^{1/2}E_\lambda(\mathbf{x})\,\boldsymbol{\varepsilon}_\lambda(a_\lambda{}^+ + a_\lambda)\,, \quad \text{(II.4a)}$$

$$\mathbf{H}(\mathbf{x},t)=-i\sum_\lambda(2\pi\hbar\omega_\lambda)^{1/2}H_\lambda(\mathbf{x})\mathbf{h}_\lambda(a_\lambda{}^+ - a_\lambda)\,. \quad \text{(II.4b)}$$

The $E_\lambda(\mathbf{x})$ and $H_\lambda(\mathbf{x})$ are normalized so that

$$(\boldsymbol{\varepsilon}_\lambda\cdot\boldsymbol{\varepsilon}_{\lambda'})\int E_\lambda(\mathbf{x})E_{\lambda'}(\mathbf{x})\,d(\mathrm{vol})=\delta_{\lambda\lambda'}\,, \quad \text{(II.5a)}$$

$$(\mathbf{h}_\lambda\cdot\mathbf{h}_{\lambda'})\int H_\lambda(\mathbf{x})H_{\lambda'}(\mathbf{x})\,d(\mathrm{vol})=\delta_{\lambda\lambda'}\,. \quad \text{(II.5b)}$$

The electric dipole moment operator for the jth atom is taken as the odd operator

$$\boldsymbol{\mu}_{\mathrm{op}}{}^j=\boldsymbol{\mu}^j(R_+{}^j+R_-{}^j)\delta(\mathbf{x}-\mathbf{x}^j)\,. \quad \text{(II.6)}$$

---

* Present address: Physics Department, University of California, Riverside, California.

[1] H. Statz and G. A. deMars in *Quantum Electronics*, edited by C. H. Townes (Columbia University Press, New York, 1960).
[2] A. A. Vuylsteke, J. Appl. Phys. 34, 1615 (1963).
[3] R. H. Dicke, Phys. Rev. 93, 99 (1954).
[4] R. H. Hellwarth, *Advances in Quantum Electronics*, edited by J. R. Singer (Columbia University Press, New York, 1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We have defined the coupling constant

$$K_\lambda{}^j=(2\pi\omega_\lambda/\hbar)^{1/2}\mu^j\cdot\varepsilon_\lambda E_\lambda(\mathbf{x}), \tag{II.7}$$

and have also defined the zero of energy as midway between the two levels of each atom so that $E_2{}^j-E_1{}^j=\hbar\Omega^j$ and $E_2{}^j+E_1{}^j=0$, where $E_2{}^j$ and $E_1{}^j$ are the energies of the upper and lower state of the jth atom, respectively.

The operator identity valid for any time-independent operator O

$$i\hbar\dot{O}=[O,\mathfrak{H}] \tag{II.8}$$

is easily worked out and gives the set

$$i\dot{R}_+{}^j=-\Omega^jR_+{}^j+2\sum_\lambda K_\lambda{}^j a_\lambda{}^+R_3{}^j, \tag{II.9a}$$

$$i\dot{R}_-{}^j=+\Omega^jR_-{}^j-2\sum_\lambda K_\lambda{}^j a_\lambda R_3{}^j, \tag{II.9b}$$

$$i\dot{R}_3{}^j=-\sum_\lambda K_\lambda{}^j(a_\lambda{}^+R_-{}^j-a_\lambda R_+{}^j), \tag{II.9c}$$

$$i\dot{a}_\lambda=\omega_\lambda a_\lambda+\sum_j K_\lambda{}^jR_-{}^j, \tag{II.9d}$$

$$i\dot{a}_\lambda{}^+=-\omega_\lambda a_\lambda{}^+-\sum_j K_\lambda{}^jR_+{}^j. \tag{II.9e}$$

Normally, expectation values of these operator identities are taken over the complete density matrix of the system (atoms+field). Taking expectation values on the right-hand side of (II.9a), (II.9b), and (II.9c) gives expectation values of products of atom and field operators, e.g., $\langle a_\lambda R_3{}^j\rangle$, etc. The problem is then essentially intractable, except in perturbation theory, for systems more complicated than a single atom interacting with a single mode. However, if we now assume that the density matrix of the system can be factored into the direct product

$$\rho=\rho_f\otimes\rho_a{}^1\otimes\rho_a{}^2\otimes\cdots\rho_a{}^N \tag{II.10}$$

of atom and field density matrices, terms such as $\langle a_\lambda R_3{}^j\rangle$ factor to the product of expectation values $\langle a_\lambda\rangle\langle R_3{}^j\rangle$. This is equivalent to phenomenological semiclassical theory, or the "S.C.F.A." (self-consistent field approximation) of Willis.[5]

It is convenient to define the slowly varying complex quantities A_λ, r^j by

$$\langle a_\lambda\rangle=A_\lambda e^{-i\omega_\lambda t}, \tag{II.11a}$$

$$\langle a_\lambda{}^+\rangle=A_\lambda{}^*e^{+i\omega_\lambda t}, \tag{II.11b}$$

$$\langle R_+{}^j\rangle=r^je^{i\Omega^jt}, \tag{II.11c}$$

$$\langle R_-{}^j\rangle=r^{j*}e^{-i\Omega^jt}. \tag{II.11d}$$

In terms of these new variables, we obtain from (II.9), after taking expectation values with respect to the factored density matrix (II.10), the set of five coupled equations

$$iA_\lambda=\sum_j K_\lambda{}^jr^{j*}e^{i(\omega_\lambda-\Omega^j)t}, \tag{II.12a}$$

$$i\dot{r}^j=2\sum_\lambda K_\lambda{}^jA_\lambda{}^*e^{i(\omega_\lambda-\Omega^j)t}R_3{}^j, \tag{II.12b}$$

$$i\dot{R}_3{}^j=-\sum_\lambda K_\lambda{}^j(A_\lambda{}^*r^{j*}e^{i(\omega_\lambda-\Omega^j)t}-A_\lambda r^je^{-i(\omega_\lambda-\Omega^j)t}), \tag{II.12c}$$

as well as their complex conjugates (where $\langle R_3{}^j\rangle$ is written simply as $R_3{}^j$).

We digress somewhat now and consider a special case of (II.12) to point out some of Dicke's results using these equations. Suppose for the moment that $\Omega^j=\Omega$ and $K_\lambda{}^j=K_\lambda$, independent of j. That is, all atoms are identical, as well as being confined in a volume whose dimensions are small compared to a wavelength $\lambda\simeq 2\pi c/\Omega$. We can then define

$$r=\sum_j r^j \tag{II.13a}$$

and

$$n=\sum_j R_3{}^j, \tag{II.13b}$$

in which special case we find the first integrals of (II.12).

$$\sum_\lambda |A_\lambda|^2+n=\text{const} \tag{II.14a}$$

and

$$|r|^2+n^2=\text{const}=R^2. \tag{II.14b}$$

(II.14a) is, of course, an expression of the conservation of energy. In analogy to Dicke's terminology, we would call (II.14b) the conservation of "cooperation." In a discussion of permanent dipole moments (spins) (II.14b) would be an expression of the conservation of total angular momentum. Here, however, we explicitly have in mind induced electric dipole moments. It is convenient to represent a radiation process geometrically in a 3-dimensional space labeling the coordinates (1,2,3) following Feynman.[6,7] According to (II.14) we can picture this simple radiation process as the motion of a vector $\mathbf{R}$ on the surface of a sphere in this space, (which, incidentally, is real space when considering permanent dipoles). The projection of $\mathbf{R}$ onto the (3) axis is the value of n, the population difference divided by two, and the projection of $\mathbf{R}$ onto the (1) axis gives $\mathrm{Re}(r)$ which is proportional to the expectation value of the induced dipole moment. The value of the induced dipole moment is zero when all atoms are exactly in their upper or lower states $n=\pm(N/2)$ and so we see that the maximum possible value of $|\mathbf{R}|=N/2$ (N is the total number of radiators), and it can have the (constant) values $(N/2)$, $(N/2)-1$, $\cdots 0$, depending on initial conditions of the problem.

[5] C. R. Willis, Bull. Am. Phys. Soc. 9, 4, 399 (1964).

[6] R. P. Feynman, F. L. Vernon, Jr., and R. W. Hellwarth, J. Appl. Phys. 28, 49 (1957).

[7] Y. H. Pao, J. Opt. Soc. Am. 52, 871 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To obtain spontaneous emission (using the same restrictive assumptions $\Omega^j=\Omega$, $K_\lambda{}^j=K_\lambda$), we simply perform a power-series expansion in powers of the coupling constant K_λ

$$A_\lambda(t)=A_\lambda{}^0(t)+K_\lambda A_\lambda{}^{(1)}(t)+\cdots,$$
$$r(t)=r^0(t)+K_\lambda r^{(1)}(t)+\cdots.$$

Inserting these into (II.12) and equating like powers of K_λ gives

$$i\dot{A}_\lambda{}^{(1)}(t)=K_\lambda r^*(0)e^{i(\omega_\lambda-\Omega)t}, \tag{II.15a}$$

$$r^0(t)=r^0(0)=r(0). \tag{II.15b}$$

$A_\lambda{}^0(t)=A_\lambda(0)=0$ for spontaneous emission. Integration of (II.9a) and insertion into (II.9c) leads to

$$\dot{n}=-|r(0)|^2\frac{d}{dt}\sum_\lambda K_\lambda{}^2\frac{\sin^2(\omega_\lambda-\Omega)t/2}{[(\omega_\lambda-\Omega)/2]^2}. \tag{II.16}$$

For spontaneous emission from the small volume into a large box of volume V, we find, in the usual way[8] (replacing the summation by an integral)

$$\dot{n}\simeq-2|r(0)|^2\rho(\Omega)K_\Omega{}^2. \tag{II.17}$$

Using

$$\rho(\Omega)=V\Omega^2/c^3 \quad\text{and}\quad \langle K_\Omega{}^2\rangle_{\text{sphere}}=2\pi\Omega\mu^2/3\hbar V \tag{II.18}$$

gives

$$\dot{n}=-|r(0)|^2/\tau=-(R^2-n^2(0))/\tau, \tag{II.19}$$

which is the desired result. Here τ^{-1} is the usual Einstein B coefficient

$$\tau^{-1}=\tfrac{4}{3}(\Omega^3\mu^2/\hbar c^3).$$

Suppose the system is started off with all atoms exactly in the upper state, $R^2=(N/2)^2$. Then, according to (II.19), nothing happens. This is a point of unstable equilibrium in the semiclassical theory. However, if $R^2=(N/2)^2$, but $N-1$ atoms are excited and one atom is in its ground state, then $n=(N/2)-1$, $\dot{n}=-(N-1)/\tau$, and the system emits "normally," that is, proportional to the number of excited atoms. But if we start off with $R^2=(N/2)^2$ and with half the atoms excited and half in their lower states, then $n(0)=0$ and $\dot{n}=-[(N/2)^2/\tau]$. The population then starts to decrease at a rate proportional to the square of the total number of atoms, and it is this situation that Dicke calls "superradiant" or "coherent" spontaneous emission. The situation is analogous to a physical pendulum which "radiates" at a very slow rate initially when aligned closely along the (3) axis but rather much more rapidly when in the (1–2) plane (when the electric field then sees a large dipole moment).

We now return to consideration of the more general set of Eqs. (II.12). Interest is in casting these into a form describing an idealized laser operating in one mode; the resulting set of 4 nonlinear coupled equation will be cast in a form which will readily yield to numerical integration.

Define the new variables m_λ, ψ_λ so that

$$m_\lambda e^{i\psi_\lambda}=\sum_j K_\lambda{}^j r^j{}^* e^{i(\omega_\lambda-\Omega^j)t}, \tag{II.20}$$

which is the (complex) dipole moment which couples into the λth mode; also define the e.m. field amplitude and phase e_λ and ϕ_λ of mode λ by

$$A_\lambda=e_\lambda e^{i\phi_\lambda}. \tag{II.21}$$

In terms of the new variables we have from (II.12)

$$\dot{e}_\lambda=-m_\lambda\sin(\phi_\lambda-\psi_\lambda), \tag{II.22a}$$

$$e_\lambda\dot{\phi}_\lambda=-m_\lambda\cos(\phi_\lambda-\psi_\lambda), \tag{II.22b}$$

$$\dot{m}_\lambda e^{i\psi_\lambda}+im_\lambda e^{i\psi_\lambda}\dot{\psi}_\lambda=\sum_{\lambda',j}2iK_\lambda{}^jK_{\lambda'}{}^j e_{\lambda'}e^{i\phi_{\lambda'}}e^{i(\omega_\lambda-\omega_{\lambda'})t}R_3{}^j+i\sum_j K_\lambda{}^j r^j{}^*(\omega_\lambda-\Omega^j)e^{i(\omega_\lambda-\Omega^j)t}. \tag{II.22c}$$

We now make the reasonable assumption that if we take an average of these equations assuming a random distribution of atomic sites $\mathbf{x}^j$, the modes will decouple due to the orthogonality condition (II.5a). That is, we now assume that

$$\langle K_\lambda{}^jK_{\lambda'}{}^j\rangle_{\text{av}}=(2\pi\omega_\lambda/\hbar V)\langle(\boldsymbol{\mu}\cdot\boldsymbol{\varepsilon}_\lambda)^2\rangle_{\text{av}}\delta_{\lambda\lambda'}\cong(2\pi\omega_\lambda/\hbar V)\mu_\lambda{}^2\delta_{\lambda\lambda'}, \tag{II.22d}$$

taking $\boldsymbol{\mu}^j\simeq\boldsymbol{\mu}$, for all j. μ_λ is the component of the dipole moment which couples into mode λ.

The second term on the right of (II.22c) would vanish if $\Omega^j=\Omega=\omega_\lambda$. It represents a dephasing of the total dipole moment due to different local environments giving rise to a spread in natural frequencies about Ω. We wish to take this term into account in a way familiar from nuclear resonance theory by saying that when $e_\lambda=0$, its effect is to cause the magnitude m_λ to decay exponentially to zero at a rate γ_λ. Then we write

$$\dot{m}_\lambda+\gamma_\lambda m_\lambda=-2\Gamma e_\lambda n\sin(\phi_\lambda-\psi_\lambda), \tag{II.23}$$

$$\dot{\psi}_\lambda m_\lambda=2\Gamma e_\lambda n\cos(\phi_\lambda-\psi_\lambda), \tag{II.24}$$

with

$$n=\sum_j R_3{}^j \tag{II.25}$$

and

$$\Gamma=2\pi\omega_\lambda\mu_\lambda{}^2/\hbar V. \tag{II.26}$$

From (II.12c), (II.20), and (II.21), we have that

$$\dot{n}=e_\lambda m_\lambda\sin(\phi_\lambda-\psi_\lambda)+\sum_{\lambda'\neq\lambda}e_{\lambda'}m_{\lambda'}\sin(\phi_{\lambda'}-\psi_{\lambda'}). \tag{II.27}$$

The sum $\sum_{\lambda'}$ runs over all those modes, except mode λ, which are at approximately frequency ω_λ but which are appreciably more lossy than the mode designated by λ. For instance, this term represents the losses of

[8] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

radiation from the sides of a lasing crystal due to spontaneous emission. The usually small stimulated emission into these modes which are characterized by high loss (free-space-like modes) will be neglected.

If we add a phenomenological mode loss rate term to (II.22a), and collect equations, we have finally the set

$$(d/dt)(e_\lambda^2)+2\beta e_\lambda^2=-m_\lambda e_\lambda \sin\theta_\lambda, \tag{II.28a}$$

$$(d/dt)m_\lambda+\gamma_\lambda m_\lambda=-2\Gamma e_\lambda n \sin\theta_\lambda, \tag{II.28b}$$

$$(d/dt)n=e_\lambda m_\lambda \sin\theta_\lambda+\sum_{\lambda'} e_{\lambda'} m_{\lambda'} \sin\theta_{\lambda'}, \tag{II.28c}$$

$$m_\lambda(d/dt)\psi_\lambda=2\Gamma e_\lambda n \cos\theta_\lambda, \tag{II.28d}$$

$$e_\lambda(d/dt)\phi_\lambda=-m_\lambda \cos\theta_\lambda, \tag{II.28e}$$

$$\theta_\lambda\equiv\phi_\lambda-\psi_\lambda. \tag{II.29}$$

From (II.28a), (II.28d), and (II.28e) we note the first integral

$$e_\lambda m_\lambda \cos\theta_\lambda=(e_\lambda m_\lambda \cos\theta_\lambda)_0 e^{-(\beta_\lambda+\gamma_\lambda)t}, \tag{II.30}$$

which states that if m_λ, e_λ, or $\cos\theta_\lambda$ were ever zero, then $\cos\theta_\lambda=0$ for all time for a nontrivial solution to exist. Otherwise, $\cos\theta_\lambda$ will approach zero at a rate $\beta_\lambda+\gamma_\lambda$. $\cos\theta_\lambda=0$ implies $\sin\theta_\lambda=\pm 1$. It must be noticed, however, that $\sin\theta_\lambda$ as a function of time takes on the two values $+1$ or -1, since by inspection of (II.28a), (II.28c) the sign of $\sin\theta_\lambda$ determines whether the atomic system is emitting or absorbing radiation.

If in (II.28b) we assume that $\dot{m}_\lambda \ll K_\lambda m_\lambda$, then we may neglect $\dot{m}_\lambda$, solve for m_λ, and insert its value into (II.28a) and (II.28c). This gives, by use of (II.30) with $\cos\theta_\lambda=0$,

$$(d/dt)e_\lambda^2+2\beta_\lambda e_\lambda^2=(2\Gamma/\gamma_\lambda)e_\lambda^2 n, \tag{II.31a}$$

$$(d/dt)n=-(2\Gamma/\gamma_\lambda)e_\lambda^2 n-(n+N/2)/\tau. \tag{II.31b}$$

The term $(n+N/2)/\tau$ has been used as an approximation for the last term of (II.28c); $n+(N/2)$ is simply the population of the excited state and τ^{-1} is the spontaneous rate. The two coupled equations constitute the well-known rate equation approach to radiation dynamics.[1] The absorption rate (per atom) for mode λ is given by $\alpha_\lambda=\Gamma/2\gamma_\lambda=\pi\omega_\lambda\mu_\lambda^2/\hbar V\gamma_\lambda$.

At the other extreme from the rate equation approximation (II.31), the completely "coherent" description of radiation into a single mode is afforded by dropping all terms in (II.19) which describe losses, that is by setting $\gamma_\lambda=\beta_\lambda=\sum e_{\lambda'} m_{\lambda'} \sin\theta_{\lambda'}=0$. In this case, by defining the angle variable Θ in the $m_\lambda/(2\Gamma)^{1/2}$, n plane as $\Theta=\tan^{-1}[(2\Gamma^{1/2})m_\lambda/n]$ we find the equation of motion for Θ

$$(d^2\Theta/dt^2)+(2\Gamma)\sin\Theta=0, \tag{II.32}$$

which is simply the equation of motion of a spherical pendulum. The initial conditions are determined from $d\Theta/dt=-(2\Gamma)^{1/2}e_{\lambda'}$. The solutions are elliptic functions and become sinusoids when $e_\lambda(0)$ gets large ($\gtrsim 10$). The case $e_\lambda(0)=0$ appropriate to spontaneous emission leads to the same general results as in Eq. (II.19) concerning the rate of spontaneouse mission as a function of N. Keeping only the mode loss rate in (II.28) results in the equation

$$(d^2\Theta/dt^2)+\beta_\lambda(d\Theta/dt)+2\Gamma \sin\Theta=0, \tag{II.33}$$

easily recognized as the equation for a damped physical pendulum.

The "perfectly coherent" process is one which describes a circle in the n, $m_\lambda/(2\Gamma)^{1/2}$ plane. The solutions to the rate equation, on the other hand, are restricted to one quadrant in this plane. Since $m_\lambda \simeq \Gamma e_\lambda n/\gamma_\lambda$, $n=0$ implies that $m_\lambda=0$, and thus the only crossing point of the axis $n=0$ is the improbable point $n=0$, $m_\lambda=0$, in the rate equation approximation.

In order to facilitate numerical integration of this set of equations, we define the new variables

$$I=e_\lambda^2/2N, \tag{II.34a}$$

$$m=m_\lambda/N(2\Gamma)^{1/2}, \tag{II.34b}$$

$$\zeta=-me_\lambda \sin\theta_\lambda/(2N)^{1/2}, \tag{II.34c}$$

$$\Delta=n/N, \tag{II.34d}$$

$$\alpha=(N\Gamma)^{1/2}. \tag{II.34e}$$

The equations (II.28) then take the form

$$\dot{I}+2\beta I=\alpha\zeta, \tag{II.35a}$$

$$dm^2/dt+2\gamma m^2=2\alpha\zeta\Delta, \tag{II.35b}$$

$$d\zeta/dt+(\beta+\gamma)\zeta=\alpha(m^2+2I\Delta), \tag{II.35c}$$

$$\dot{\Delta}=-\alpha\zeta-[(\Delta+\tfrac{1}{2})/\tau]-P(\Delta-\tfrac{1}{2}), \tag{II.35d}$$

where we have suppressed the subscript λ, written the term $\sum_{\lambda'} e_{\lambda'} m_{\lambda'} \sin\theta_{\lambda'}$ as $(\Delta+\frac{1}{2})/\tau$, and added a term to the equation for the rate of change of the population difference Δ (normalized to $2N$) to describe the effect of a "pump" which takes atoms from the ground state and puts them into the upper state. We are thus assuming an infinitely fast transition from the pumping band to the upper laser level, as well as an optically thin material so that the pumping is uniform throughout the volume V. We can allow P (the pumping rate) as well as β, the cavity mode loss rate, to be functions of time; the numerical solutions below, however, only include the case of a time varying β.

III. RESULTS OF NUMERICAL INTEGRATION

The set of Eqs. (II.35) were integrated on a high-speed digital computer for several values of a (constant) pump rate P, and the coupling parameter α, and the dipole moment dephasing rate γ. The function β is taken to be a "Fermi function" of time

$$\beta(t)=[(\beta_i-\beta_f)/(1+e^{(t-t_0)/\Delta t})]+\beta_f, \tag{III.1}$$

and will represent approximately the "Q switching" by a Kerr cell or similar apparatus. t_0 will be chosen as a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

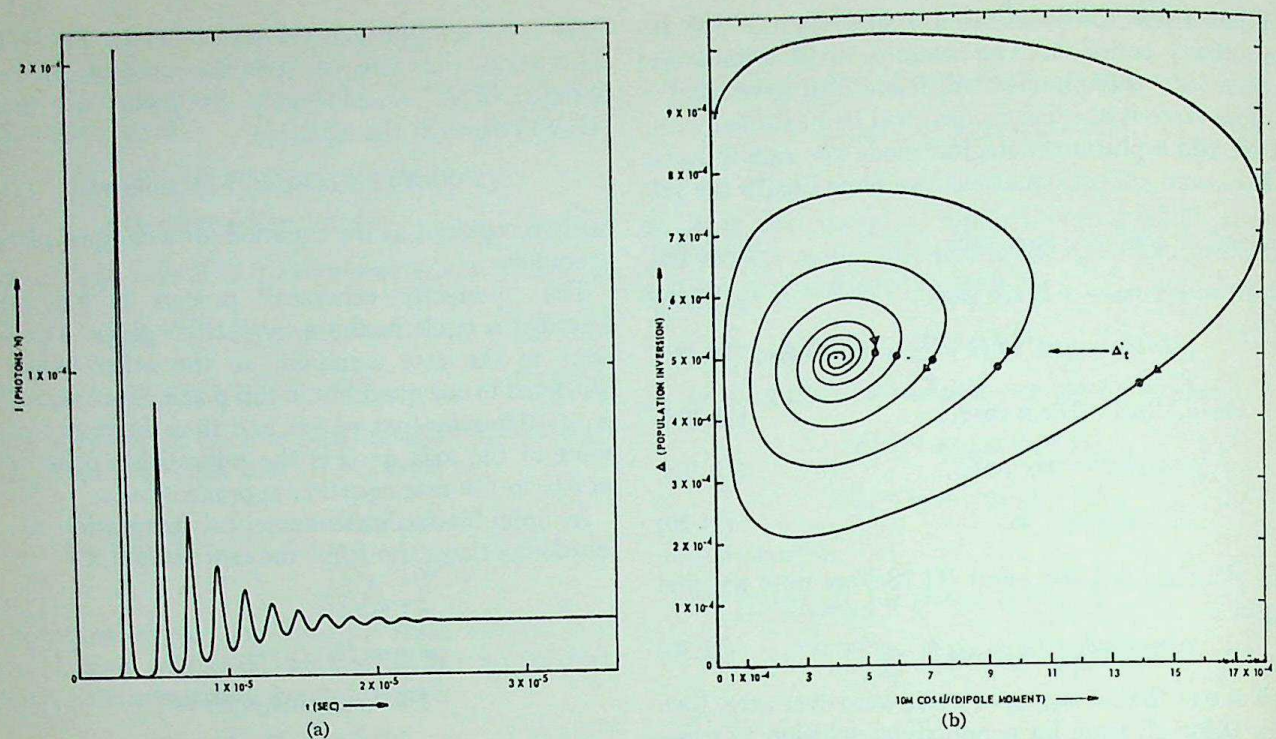


FIG. 1. Laser behavior with constant pump and constant cavity loss rate under conditions which would be well approximated by the rate equation approach. Parameter values are $\beta=10^7$, $\gamma=10^8$, $P=10^3$, $\alpha=10^9$ and $\tau^{-1}=3\times10^2$, all in $\sec^{-1}$. $t=0$ is chosen from computed data near the start of significant behavior. (a) Intensity dependence on time. (b) Dipole moment versus population inversion.

time somewhere just before the populations Δ reach a threshold determined by β_i.

Figure 1 shows the "relaxation oscillations" for a laser operating under conditions of ordinary power output.[9] Both the intensity I as a function of time as well as the corresponding phase diagram of $(m\cos\psi)$ versus Δ is shown for each case. The dots in the phase diagrams correspond in time to the peaks in the intensity diagram. In Fig. 1(b), the entire phase diagram is seen to lie in the first quadrant and the solution is very well predicted by the rate equations, since the rate γ is much greater than the fractional rate of change of m. The intensity I is shown in units of photons/N, where N is the effective number of atoms participating. By changing the cavity loss rate to a larger value we see in Fig. 2, as expected, that the

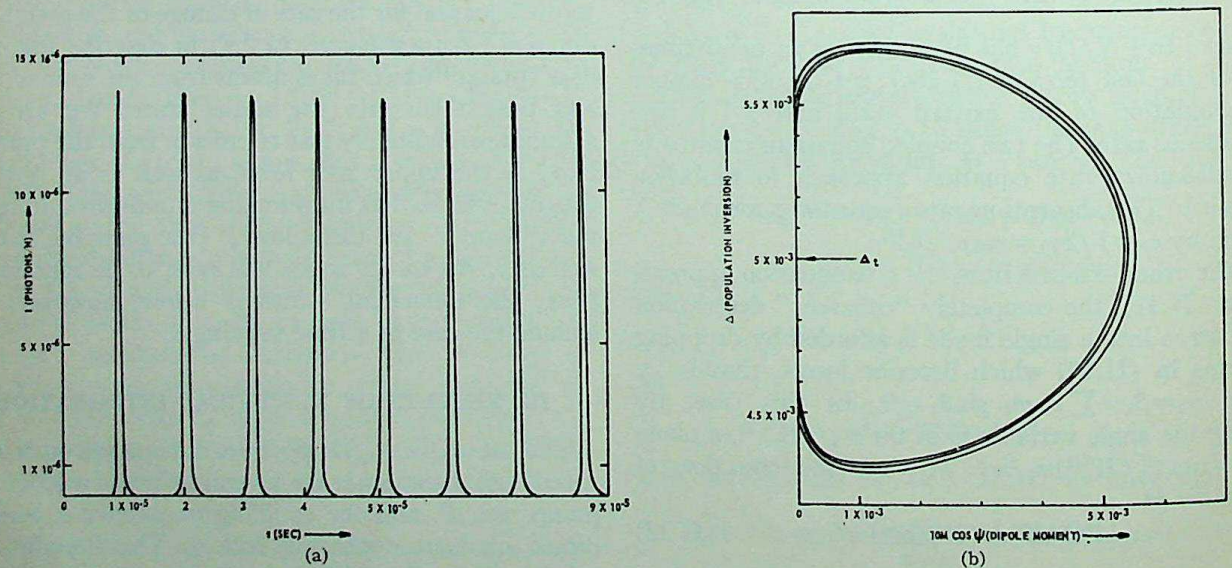


FIG. 2. Same as Fig. 1, except for higher cavity loss rate, $\beta=10^8$ $\sec^{-1}$.

[9] For all the following figures, the quantity $(e_\lambda m_\lambda \cos\theta_\lambda)_0$ was taken as zero. [See Eq. (II.30).] Moreover, for convenience in graphing, the same discontinuous properties are attributed to ψ and ϕ as are exhibited by $\theta=\phi-\psi$. That is, $\sin\phi$ and $\sin\psi$ as functions of time are made to alternate between the values $+1$ and -1 in such a manner that $\sin\theta(t)$ takes on the correct values and the functions $m\cos\psi$ and $e\sin\phi$ are continuous.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

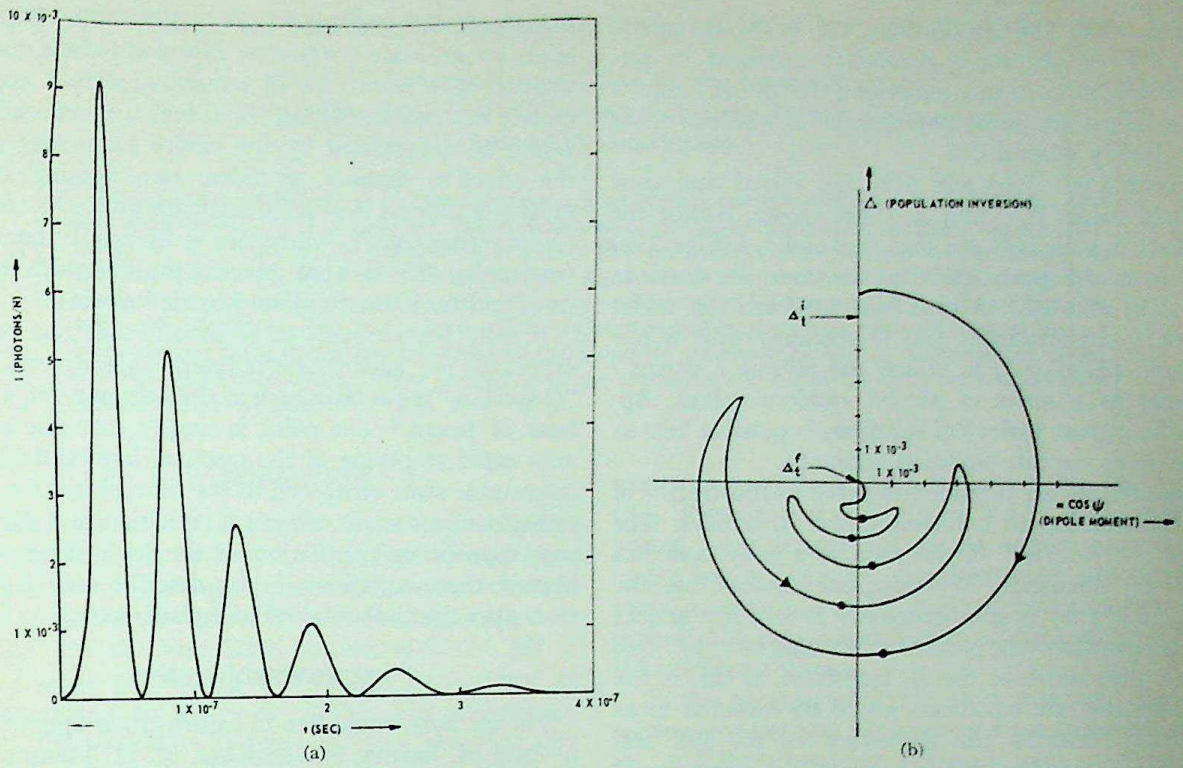


FIG. 3. Same as Fig. 1, except for a larger pump $P=10^4$ sec^{-1}, a smaller dephasing rate $\gamma=10^7$ sec^{-1}, and the addition of Q spoiling with cavity loss rate initially $\beta_i=10^9$ sec^{-1} changing to a final cavity loss rate of $\beta_f=10^7$ sec^{-1} in a time $\Delta t=2.5.10^{-8}$ sec.

spiking is almost continuous; clearly the radiation almost completely escapes from the cavity before the pump can again bring the population over threshold, $\Delta_t=\beta\gamma/2\alpha^2$, and the process simply repeats itself. (The parameter τ, the spontaneous emission time has been taken as 3×10^{-3} sec in all computations.) Figure 3 shows a case of "Q switching" [Eq. (III.1)], where parameters are chosen to exhibit a rather extreme situation (in present day practice) of very low γ and β.

The second, third, etc., spikes in Fig. 3(a) are seen

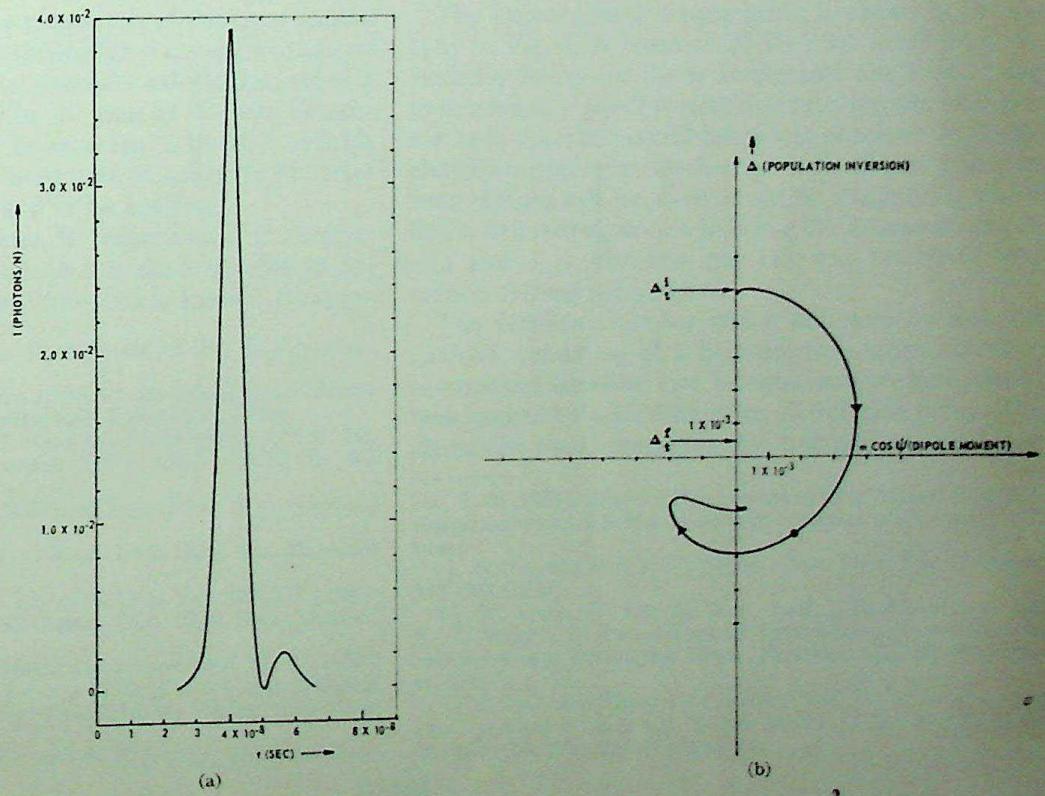


FIG. 4. Same as Fig. 3, except that final cavity loss rate $\beta_f=10^8$ sec^{-1} and Q spoiling occurs over time $\Delta t=5.10^{-9}$ sec.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to occur when the populations are below inversion, $\Delta<0$. This is strictly a possibility inherent in this analysis and not in a rate equation description. Figure 4 again exhibits the same phenomena; the second spike occurs totally when $\Delta<0$.

In both Figs. 3(b) and 4(b) the initial and final threshold values $\Delta_t^{(i)}$ and $\Delta_t^{(f)}$ are shown. Notice the difference in scale between Figs. 1, 2 and 3, 4. Figs. 1(b) and 2(b) would appear as diffuse dots near the Δ axis in Figs. 3(b) and 4(b). A case has also been run under "normal" circumstances (no Q spoiling) with a low threshold ($\Delta_t\approx 0.001$) in which the output I versus t appeared as a series of almost random spikes. Apparently, a great variety of solutions is possible here as compared to the rate equation approach.

These results are intended to serve as illustrative of some possibilities in coherent device dynamics. The case of a "two-level" device akin to a ruby laser has guided the discussion. However, one expects that the laser line in ruby is homogeneously broadened, and at room temperatures the parameter γ would be $\approx 10^{11}$ cps, the full time width of the R_1 transition. If this is the case then the rate equations would be expected to as adequately describe the dynamics as the full set, (II.35) and secondary pulses as described above would not be possible unless one operated as a much lower temperature, $\lesssim 70°$K. However, if the laser line is inhomogeneously broadened even at room temperature, (which appears unlikely in ruby), then such "hopping" might occur, since γ would be expected to be a good bit smaller. The above set of equations should approximately describe the dynamics of such a system where γ would be determined by the cavity parameter β and the effective number of atoms participating in the radiation process into a single mode would be roughly $(N)\text{eff}\approx(\Delta\omega_c/\Delta\omega)N$, where $\Delta\omega_c$ is the cavity linewidth (presuming this is what governs the linewidth in this case) and $\Delta\omega$ is the full laser atomic linewidth.

It may be pertinent to point out that one cannot expect to see such "superradiant" hopping effects by "Q spoiling" a gas laser such as the helium-neon, 4-level laser of Javan.[10] The point is simply that due to the very rapid emptying of the terminal laser state into a metastable state compared to the time width of a giant pulse, one can never achieve a condition of $\Delta<0$ with any appreciable population in the lower laser state; that is, these equations do not describe such a device even over the time interval of a giant pulse.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Paul M. Sutton, Sylvan Katz, M. H. Johnson, and M. B. White for valuable discussions.

[10] A. Javan, W. R. Bennett, Jr., and D. R. Herriot, Phys. Rev. Letters 6, 106 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Collisional Electron Detachment from 20-MeV D^0 and D^- Ions*

KLAUS H. BERKNER, SELIG N. KAPLAN, AND ROBERT V. PYLE
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received 27 January 1964)

Cross sections for collisional electron detachment from 20-MeV deuterium atoms and negative ions have been measured in H_2, He, N_2, and Ar. The cross sections σ_{01} per target molecule for the ionization of neutral atoms are 2.2, 1.6, 20, and 36, respectively, in units of 10^{-18} cm^2. Results for single-electron detachment from D^- ions, σ_{-10}, are 7.5, 4.7, 58, and 75, respectively. All values have an uncertainty of $\pm 10\%$, chiefly from absolute-pressure calibration uncertainties. The experimental values for atoms in H_2 and He are in agreement with calculations that use the Born and free-collision approximations. The measured cross sections for single-electron detachment from negative ions in H_2 and He are about one-half of the calculated values. The two-electron detachment cross sections, σ_{-11}, are approximately 4% of the single-detachment cross sections, σ_{-10}, in all gases.

I. INTRODUCTION

THE cross sections for electron loss from fast hydrogen atoms in hydrogen and helium targets have been calculated by using three models: the semiclassical model of Bohr,[1] the "free-collision" (impulse) approximation,[2] and the first Born approximation.[3,4] All of these models predict an E^{-1} energy dependence at high energies (i.e., where $Ze^2/\hbar v \ll 1$). For a hydrogen target, this condition implies energies above a few hundred kilovolts. In this asymptotic region, the Born and free-collision approximations give numerical values that agree with each other and with experiment[5] to within 15%.

For single-electron detachment from fast H^- ions, the evaluation of the cross section by the Born approximation has been considered too cumbersome to carry out with the best available H^- wave function, and the result is so sensitive to the choice of other approximate wave functions that the magnitude of the cross section is very uncertain.[6] The free-collision approximation, on the other hand, permits a simpler calculation; since it has given good results in the case of electron detachment from H^0 it might be expected to be fairly reliable in the H^- case. It has been used to calculate H^- cross sections in H and He[2] and in N_2 and O_2.[7]

We know of only three H^- experiments at energies that are sufficiently high to test the free-collision approximation. Rose *et al.*[8] have made gas-cell measurements at energies up to 1.8 MeV in H_2, O_2, Ar, and CO_2. Fremlin and Spiers[9] and Verba *et al.*[10] have reported measurements in air at higher energies (but under less well-known conditions) by observing the radial attenuation of H^- internal cyclotron beams. In general, the experimental cross sections are well below the calculated values.

In the present work we have measured the cross sections for electron detachment from hydrogen atoms, σ_{01}, and from negative hydrogen ions, σ_{-10} and σ_{-11}, in a single high-energy experiment under well-controlled conditions.

We note that the magnitudes of these cross sections and the way they extrapolate to higher energies are of interest in certain high-energy-accelerator design considerations.[7,9]

II. EXPERIMENTAL ARRANGEMENT

The experimental arrangement is shown schematically in Fig. 1. A beam of 20-MeV D^- ions[11] from the Berkeley heavy-ion linear accelerator was bent 15 deg to remove any possible contaminants. For the measurements of σ_{01}, the neutral beam was obtained by single-electron stripping in the first gas cell; all ions emerging from the gas cell were swept out by magnet $S1$, which had a field strength of 4 kG. For the measurements of σ_{-10} and σ_{-11}, the first gas cell was evacuated and magnet $S1$ was set at 0 kG.

The collision chamber was a differentially pumped gas cell, consisting of a high-pressure target chamber sandwiched between two intermediate-pressure chambers backed by a 1500-liter/sec oil diffusion pump. The chambers were connected by 6-mm-diam 5-cm-long

* This work done under the auspices of the U. S. Atomic Energy Commission.

[1] N. Bohr, *The Penetration of Atomic Particles Through Matter* (I Kommission Hos E. Munksgaard, København, 1948).

[2] I. S. Dmitriev and V. S. Nikolaev, Zh. Eksperim. i Teor. Fiz. 44, 660 (1963) [English transl.: Soviet Phys.—JETP 17, 447 (1963)].

[3] D. R. Bates, and A. Williams, Proc. Phys. Soc. (London) A70, 306 (1957).

[4] D. R. Bates and G. W. Griffing, Proc. Phys. Soc. (London) A68, 90 (1955).

[5] C. F. Barnett and H. K. Reynolds, Phys. Rev. 109, 355 (1958).

[6] M. R. McDowell and G. Peach, Proc. Phys. Soc. (London) 74, 463 (1959).

[7] B. T. Wright, Arch. Math. Naturvidenskab. 53, 9 (1957). An asymptotic expression for σ_{-10} in Ar is also given; it should be applicable only at energies considerably higher than those discussed here.

[8] P. H. Rose, R. J. Connor, and R. P. Bastide, Bull. Am. Phys. Soc. 3, 40 (1958). Numerical values taken from Allison (Ref. 12). Estimated errors are from a private communication from P. H. Rose.

[9] J. H. Fremlin and V. M. Spiers, Proc. Phys. Soc. (London) A68, 398 (1955).

[10] J. W. Verba, W. Kündig, A. C. Paul, J. R. Richardson, and B. T. Wright, in Proceedings of International Conference on Sector Focused Cyclotrons Meson Factories, CERN, 1963 (unpublished), p. 95.

[11] We used deuterium ions because the Berkeley heavy-ion linear accelerator does not operate stably at the low gradients necessary for hydrogen acceleration.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

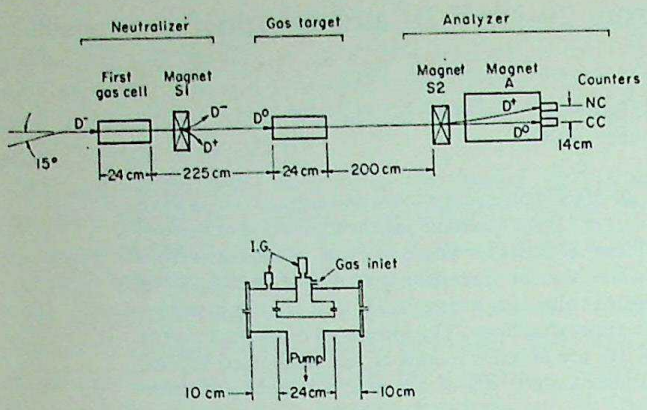


FIG. 1. The experimental arrangement.

tubes (inset Fig. 1). With this cell a base pressure of 5×10^{-6} Torr could be maintained in the drift sections when the pressure in the target chamber was 5×10^{-2} Torr. The effective length of the target chamber was $L=24\pm1$ cm. Integrated target densities ranged from $NL=0.5$ to 30×10^{15} atoms/cm^2 for Ar and N_2, and $NL=2$ to 200×10^{15} atoms/cm^2 for He and H_2. The pressure in the target chamber was monitored by a Westinghouse type-7676 high-pressure ionization gauge (Schulz-Phelps gauge) which was cross calibrated with three liquid-nitrogen-trapped McLeod gauges. Fluctuations of the calibration from day to day and uncertainty of the absolute accuracy of the McLeod gauges indicate an uncertainty of $\pm10\%$ in the pressure.

The ions in the beam emerging from the gas target were deflected by magnet $S2$ (3 kG) and steered into the detector CC by magnet A($\approx$8 kG), the neutrals were detected by NC. The detectors NC and CC each consisted of a 5-cm-diam plastic scintillator and photomultiplier. By means of collimating slits it was determined that the ions reaching the detector were well localized in a 5-mm-diam beam spot, which was small compared to the size of the detector. If two species of ions were to be detected, magnet A was used to steer first one and then the other into the detector CC, while the neutral counter was used as a monitor.

III. RESULTS AND ANALYSIS

Cross sections were calculated at a number of pressures for each target gas according to the following formulas. (We neglect cross sections for electron capture by the fast ions; these are at least three orders of magnitude smaller than the electron loss values considered here.[12])

If n_0 is the number of hydrogen atoms at a depth π in the target, σ_{01} is the cross section for electron detachment, $\Pi\equiv NL$ is the target thickness (atoms/cm^2), and n_+ is the number of positive ions produced in the thickness π, then we have

$$dn_0/d\pi=-n_0\sigma_{01}=-dn_+/d\pi, \tag{1}$$

with the solution

$$\sigma_{01}=\frac{1}{\Pi}\ln\left[1+\frac{n_+(\Pi)}{n_0(\Pi)}\right]. \tag{2}$$

Similarly, the disappearance rate for negative hydrogen ions due to loss of one electron (σ_{-10}) and two electrons (σ_{-11}) in a single collision is given by

$$dn_-/d\pi=-n_-(\sigma_{-10}+\sigma_{-11}), \tag{3}$$

with the solution

$$\sigma_{-10}+\sigma_{-11}=\frac{1}{\Pi}\ln\left[1+\frac{n_+(\Pi)+n_0(\Pi)}{n_-(\Pi)}\right]. \tag{4}$$

Using the additional relation

$$dn_0/d\pi=n_-\sigma_{-10}-n_0\sigma_{01}, \tag{5}$$

we obtain

$$\sigma_{-10}=\frac{n_0(\Pi)}{n_-(\Pi)}\left[\frac{\sigma_{01}-(\sigma_{-10}+\sigma_{-11})}{1-\exp\Pi(\sigma_{-10}+\sigma_{-11}-\sigma_{01})}\right]. \tag{6}$$

Equations (4) and (6) were solved for σ_{-10} and σ_{-11}, and Eq. (2) was solved for σ_{01}. Uncertainties in the pressure, cell length, background correction, and counting statistics were propagated through the above equations to obtain the uncertainties in the cross sections.

At sufficiently low pressures, the above equations and the uncertainties in the results can be approximated by relatively simple expressions that demonstrate the correct limiting behavior.

The σ_{01} calculations quoted in the text have been carried out for fast hydrogen atoms in the ground state. If an appreciable fraction of the atoms of our beam were in excited states, and if the variation of σ_{01} with quantum level, n, were sufficiently large, then the experimental values could be considerably greater than the ground-state cross sections. This situation would show up as a decrease in the measured cross sections with increasing gas-cell pressure, closely approaching the ground-state value at the higher pressures. Unfortunately, we did not have time to take σ_{01} over a large pressure range (the maximum was a factor of 10 in H_2), but no variation of the kind mentioned above occurred in any of the gases.

It is known[13,14] that perhaps 15% of the neutral atoms produced by charge exchange or H_2^+ breakup are excited to $n\gtrsim3$; the fraction in each state is $f\approx(2-3)n^{-3}$. (The $n=2$ states completely decay in our geometry.) Preliminary measurements by us indi-

[12] S. K. Allison, Rev. Mod. Phys. **30**, 1137 (1958).

[13] A. C. Riviere and D. R. Sweetman, in *Comptes Rendus de la VIe Conférence Internationale sur les Phénomènes d'Ionisation dans les Gaz* (Paris 1963), Vol. 1, 105.

[14] K. H. Berkner, J. R. Hiskes, S. N. Kaplan, G. A. Paulikas, and R. V. Pyle, in *Proceedings of the Third International Conference on the Physics of Electronic and Atomic Collisions, University College, London, 1963* (North-Holland Publishing Company, Amsterdam, in press).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

cate that the excited-state populations are an order of magnitude smaller when the atoms are produced by electron detachment from negative hydrogen ions, as in this experiment. There is also evidence[14,15] that the detachment cross section varies more slowly than n. We therefore conclude that the measured cross section is appropriate to the ground state. This question does not arise for the negative-hydrogen-detachment cross sections.

Results for nitrogen are shown in Fig. 2 in order to illustrate the degree of internal consistency. Results for other targets are of similar quality.

The weighted averages of the various cross-section measurements are given in Table I. To aid in the discussion, they are also plotted in Fig. 3 together with lower energy data from other experiments and theoretical estimates. In this figure, and in the following discussion, the ions are treated as if they were 10-MeV hydrogen rather than 20-MeV deuterium particles.

IV. DISCUSSION AND CONCLUSIONS

A. Electron Detachment from Neutral Atoms, σ_{01}

Our measured cross sections for electron detachment from neutral hydrogen in H_2 and He are in excellent agreement with the theoretical predictions of the free-collision[2] and Born approximations[3,4] and are consistent with an E^{-1} extrapolation of the lower energy results of Barnett and Reynolds[5] (Fig. 3). For the heavier targets, N_2 and Ar, such cross-section calculations have not been reported. A linear extrapolation to the results of Barnett and Reynolds indicates that certainly in Ar, and probably also in N_2, the cross section falls off less rapidly than the E^{-1} predicted for and observed in the lighter targets. The semiclassical theory of Bohr,[1] which also predicts the E^{-1} energy dependence in light elements (although it considerably overestimates the magnitudes of the cross sections), predicts a weaker energy dependence, approximately proportional to $E^{-1/2}$ for medium-Z elements. For reference, lines corresponding to E^{-1}, $E^{-2/3}$, and $E^{-1/2}$ variations are drawn through our σ_{01} point in Ar, but these are not intended to suggest that the asymptotic energy region has been reached.

TABLE I. Measured cross sections per target molecule for 20-MeV deuterium (in units of 10^{-18} cm^2).

	σ_{01} ±10%	$\sigma_{-11}+\sigma_{-10}$ ±10%	σ_{-10} ±10%	σ_{-11} ±25%	$\frac{\sigma_{-11}}{\sigma_{-10}}$	$\frac{\sigma_{01}}{\sigma_{-10}}$
H_2	2.2	7.8	7.5	0.29	0.038	0.31
He	1.6	4.9	4.7	0.14	0.029	0.34
N_2	20	61	58	2.7	0.046	0.34
Ar	36	78	75	2.5	0.033	0.48

[15] A. C. Riviere and D. R. Sweetman, in *Proceedings of the Third International Conference on the Physics of Electronic and Atomic Collisions, University College, London, 1963* (North-Holland Publishing Company, Amsterdam, to be published).

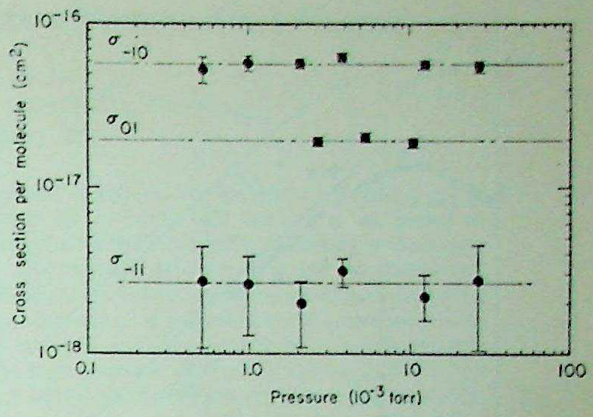


FIG. 2. Cross-section measurements in N_2 versus pressure.

B. Single-Electron Detachment from Negative Hydrogen Ions, σ_{-10}

The results for single-electron loss from negative hydrogen ions, σ_{-10}, seem to show the same energy dependence as the σ_{01} cross sections at high energies (Fig. 3). The Born-approximation calculation of McDowell and Peach[6] for a hydrogen target predicts a cross section that is $2\frac{1}{2}$ times our experimental value. Moreover, in the 1–10-MeV energy range their calculated cross section has not yet reached the asymptotic dependence and thus varies much more slowly than E^{-1}. (The line marked M-P in Fig. 3 is their asymptotic solution.) As mentioned in the introduction, this calculation is extremely sensitive to the choice of the approximation to the H^- wave function. Because of this, they state that their results should not be in error by more than a factor of 5 at high energies.

The free-collision approximation, which requires only a knowledge of the ionization potential, also gives results that are higher than the experimental values for H_2,[2] He,[2] and N_2.[7]

C. The Two-Electron Detachment Cross Sections, σ_{-11}

The cross section for the loss of both electrons from an H^- ion, σ_{-11}, is approximately 4% of the one-electron detachment cross section, σ_{-10}, at 10 MeV in all of the target gases. This proportion is similar to that observed by Fogel *et al.* at energies of 40 keV or less.[16] We know of no theoretical predictions for the σ_{-11} cross sections; the problem is currently under consideration by McDowell.[17]

[16] Ia. M. Fogel, V. A. Ankudinov, and R. E. Slabospitskii, Zh. Eksperim. i Teor. Fiz. 32, 453 (1957) [English transl.: Soviet Phys.—JETP 5, 382 (1957)].

[17] M. R. McDowell, Royal Holloway College, England (private communication).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

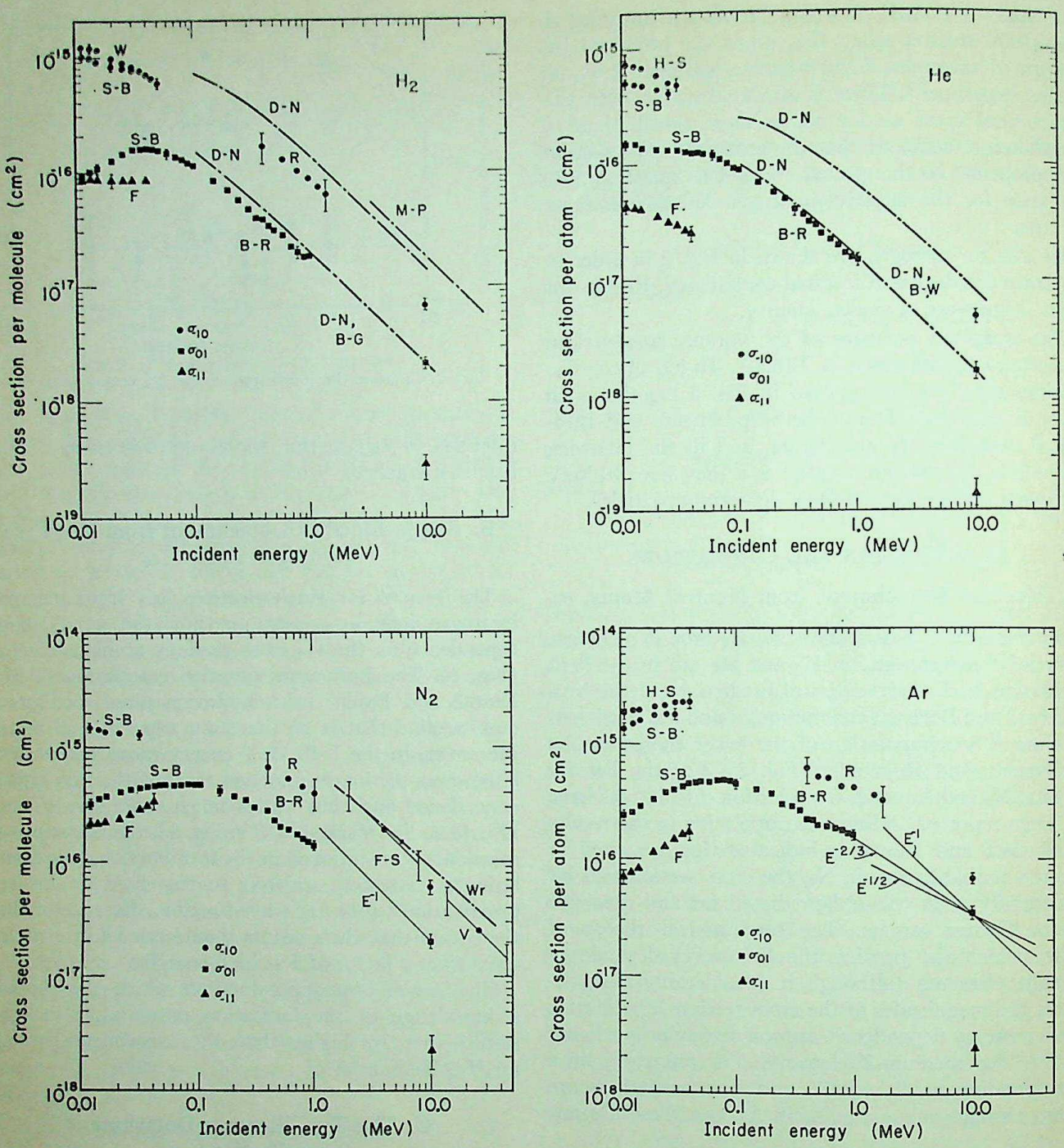


FIG. 3. Energy dependence of the hydrogen cross sections. ●:σ_{-10}; ■:σ_{01}; ▲:σ_{-11}. The "10-MeV" points are the 20-MeV deuterium data of this paper, and the others are from Allison's review article (Ref. 12), which includes measurements by Barnett and Reynolds (B-R), Fogel *et al.* (F), Hasted and Stedeford (H-S), Rose *et al.* (R), Stier and Barnett (S-B), and Whittier (W). The solid lines are internal cyclotron-beam measurements by Fremlin and Spiers (F-S) (Ref. 9) and Verba *et al.* (V) (Ref. 10). The broken lines give Born-approximation calculations by Bates and Williams (B-W) (Ref. 3), Bates and Griffing (B-G) (Ref. 4), and McDowell and Peach (M-P) (Ref. 6), and free-collision-approximation calculations by Dmitriev and Nikolaev (D-N) (Ref. 2) and Wright (Wr) (Ref. 7).

D. Summary

The measured σ_{01} cross sections are in good agreement with theory for H_2 and He target gases; calculations are not available in higher Z materials. The σ_{-10} cross sections are not in good agreement with theory. As yet, no calculations of σ_{-11} have been reported.

Measurements at considerably higher energies are necessary before the asymptotic nature of many of the cross sections can be established.

ACKNOWLEDGMENTS

We are grateful to Dr. Chester M. Van Atta for supporting and encouraging this research. One of us (SNK) thanks Dr. Burton J. Moyer for the support and interest that enabled him to participate. We also thank Dr. Marvin H. Mittleman for helpful discussions. The help of Dr. George A. Paulikas, Dr. Henry F. Rugge, Vincent J. Honey, and J. Warren Stearns with the experiment is greatly appreciated.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Hysteresis in Stability Conditions of Electron-Hole Plasma

BETSY ANCKER-JOHNSON
Boeing Scientific Research Laboratories, Seattle, Washington
(Received 13 January 1964)

The investigation of the hysteresis occurring in the threshold conditions for the helical instability oscillations in electron-hole plasma has been extended. The hysteresis in electric field strength E can exceed 45 V/cm and 50% of the applied E at threshold. By determining the stability-instability boundary in the p-type InSb as a function of the parallel (or antiparallel) electric and magnetic field strengths and also the plasma density, the magnetic field induced by the formation of the helical density perturbation is deduced. Resulting displaced loop B-H curves are presented. The magnitude of the induction enhancement B_{hys} can be as large as 165 Oe and 55% of the applied B at threshold. The extent of applied magnetic field (or electric field strength) over which the loops can occur is limited at the high magnetic field end (typically $\lesssim$600 G) by vanishingly small plasma density and at the other end ($\gtrsim$280 G) by the occurrence of magnetic pinching. The loops are largest at the low magnetic field end. The range of loop energy, defined as the product of E_{hys} and B_{hys}, increases with decreasing plasma cross-sectional area until saturation occurs at a plasma radius of $\sim 3\times10^{-2}$ cm. The full range from largest to smallest loop in any one sample is achieved by a relatively small variation in the "input energies," i.e., the product of the magnetic and electric field strengths at threshold.

I. INTRODUCTION

IT is well known that the stability of a plasma may be enhanced by the application of a longitudinal magnetic field.[1] Early stability and confinement experiments also showed the existence of a definite limit to the improvement so obtainable. Indeed, too large a magnetic field causes the onset of oscillations in the plasma and its rapid radial diffusion to the container boundaries. The helical instability theory first proposed by Kadomtsev and Nedospasov[2] accounts for these occurrences and describes the boundary between stability and instability in terms of the applied, parallel electric and magnetic field strengths. Johnson and Jerde[3] supplied a firm mathematical base for the theory and improved its predictions.

Almost concurrently with these electron-ion plasma experiments,[1] Ivanov and Ryvkin[4] reported the occurrence of current oscillations in germanium immersed in a longitudinal magnetic field. Larrabee and Steele[5] made a detailed study of this phenomenon and named it the oscillistor. They demonstrated that the oscillations occur only if excess carriers, i.e., an electron-hole plasma, is present in the semiconductor. Glicksman[6] recognized the applicability of Kadomtsev and Nedospasov's theory to this phenomenon and adapted it to electron-hole plasmas in an insulator. Holter[7] used Johnson and Jerde's[3] approach to derive again explicit expressions for the threshold frequency and electric field strength as functions of magnetic field strength, some parameters of the plasma and its container (the semiconductor), this time applicable also to both extrinsic and intrinsic semiconductors as well as to insulators. His comparison with Ancker-Johnson's[8] measurements of the threshold conditions showed good agreement. Other data[9,10] on the threshold conditions which do not include the densities of the plasmas, a parameter of the theory, are compatible with theory.

Recently, hysteresis in the conditions necessary for the production of the helical instability oscillations in electron-hole plasmas was reported along with the suggestion of its practical application to memory devices.[11] This hysteresis differs from the familiar type possessed, for example, by ferrites, in that it exhibits two loops which are displaced from the origin of the B–H curve, as illustrated by Fig. 1(a). A plasma is

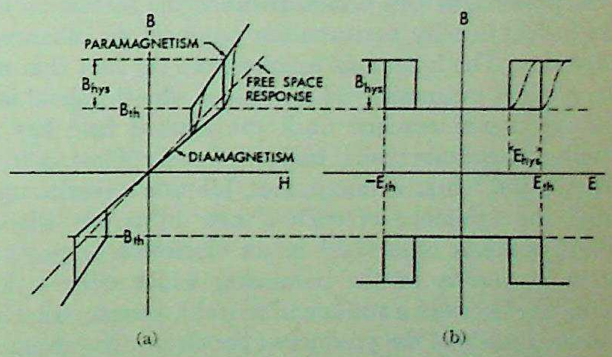


FIG. 1. Hysteresis loop diagrams for helical instability oscillations: (a) B–H curves; (b) B–E curves. The dot-dash line in each first quadrant shows another possible loop shape.

[1] B. Lehnert, in *Proceedings of the Second United Nations International Conference on Peaceful Uses of Atomic Energy, Geneva* (United Nations, Geneva, 1958), p. 146.
[2] B. B. Kadomtsev and A. V. Nedospasov, J. Nucl. Energy 1, 230 (1960); see F. C. Hoh, Rev. Mod. Phys. 34, 267 (1962) for a review of experiments and theories of the helical instability in electron-ion plasmas.
[3] R. R. Johnson and D. A. Jerde, Phys. Fluids 5, 988 (1962).
[4] I. L. Ivanov and S. M. Ryvkin, Zh. Techn. Fiz. 28, 774 (1958) [English transl: Soviet Phys.—Tech. Phys. 3, 722 (1958)].
[5] R. D. Larrabee and M. C. Steele, J. Appl. Phys. 31, 1519 (1960); R. D. Larrabee, J. Appl. Phys. 34, 880 (1963); see also, J. Bok and R. Veilex, Compt. Rend. 248, 2300 (1958).
[6] M. Glicksman, Phys. Rev. 124, 1655 (1961).
[7] O. Holter, Phys. Rev. 129, 2548 (1963); for more details see Boeing Scientific Research Labs Document D1-82-0198, September 1962 (unpublished), or Årbok Univ. i Bergen, Mat.-Nat., Ser. 8, 3 (1963).
[8] B. Ancker-Johnson, in *Proceedings of the International Conference on Physics of Semiconductors, Exeter* (The Institute of Physics and the Physical Society, London, 1962), p. 131.
[9] (a) T. Misawa and T. Yamada, Japan. J. Appl. Phys. 2, 19 (1963); (b) T. Misawa, Japan. J. Appl. Phys. 1, 67 and 131 (1962).
[10] F. Okamoto, T. Koike, and S. Tosima, J. Phys. Soc. Japan 17, 804 (1962).
[11] B. Ancker-Johnson, Appl. Phys. Letters 3, 104 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Sample properties and hysteresis data. See Fig. 3 for dimensions.

Sample No.	Initial hole density (cm^{-3})	Hole mobility (cm^2/V-sec)	E_{th}(V/cm)	B_{th}(G)	ξ	E_{hys}(V/cm)	B_{hys} deduced from Experimental ξ boundary	B_{hys} deduced from Constant ξ boundary
2D-2	4.9×10^{14}	8600	99.5	282	0.53[a]	35	65	20
			86.5	302.5	0.24[a]	26	65	18
			75	317	0.15[a]	17	62	13
			69	337	0.08_5	13	48	9
			63.5	355	0.04_5	10	48	8
			56.5	385	0.02_3	9	59	7
			51.5	417	0.01_2	6	38	6
2D2-7	5.0×10^{14}	8300	89.5	299	0.12_2[a]	48.5	165.5	61
			74	325	0.07_4[a]	34	144	68
			66	365	0.05_6	29	114	47
			63	376	0.04_9	21	83	36
			60.5	390	0.04_2	14.5	54	20
			52	421	0.02_8	10	38	11
2D2-8	4.3×10^{14}	7200	164	345	0.08_7	40	93	31
			159	370	0.07_8	42	91	26
			143	394	0.05_8	32	88	18
			127	430	0.03_2	24	80	26
			118	460	0.03_1	19	64	21
			110	485	0.02_2	14.5	51	15
			103	510	0.01_8	9.5	34	6
2D2-2	3.7×10^{14}	7000	85.5	390	0.11_3	22.5	69	37
			79	410	0.10_0	18	59	28
			73	430	0.08_5	15	55.5	24
			66	450	0.07_2	13	59.5	...
			60	471	0.06_1	11.5	70.5	27
			57	490	0.05_5	10.5	77	...
			53	510	0.04_8	9	82	30
			50	530	0.04_2	8	87	30
			48	550	0.03_9	7	80	28
			46	570	0.03_3	6	70	24
			44.5	590	0.03_2	5	53	20
			42.5	610	0.02_7	3	36	12

[a] Obtained by extrapolation of the I–V characteristic after the onset of current pinching.

normally diamagnetic but at the onset of the helical instability it becomes paramagnetic as experimentally demonstrated in an electron-ion plasma by Johnson.[12] This paramagnetism occurs because the formation of the helical density perturbation produces an enhanced induction. The hysteresis follows from the fact that as the applied magnetic field is reduced, the threshold induction is not reached until the applied field has a smaller magnitude than it had at the onset of instability. Drummond[13] first deduced that Johnson's results implied the existence of such a new hysteresis effect. The hysteresis observed[11] in an electron-hole plasma is an hysteresis in the parameter which defines the boundary between a stable and unstable plasma, namely the magnitude of the produce of the electric and magnetic fields, as illustrated in Fig. 1(b). The present paper extends these hysteresis measurements to the determination of the induction enhancement caused by the helical instability and, hence, B–H curves of the type illustrated in Fig. 1(a) are presented.

[12] R. R. Johnson, in *Proceedings of the Sixth International Conference on Ionization Phenomena in Gases* (Paris, 1963), Vol. I, p. 413.

[13] J. E. Drummond, Boeing Scientific Research Laboratories Progress Review, First Six Months 1963 (unpublished), p. 95.

II. EXPERIMENT

The circuit is the same as that employed previously.[11] It consists simply of a triangular waveform delivered to a single crystal parallelepiped of p-type InSb at 77°K in the presence of a magnetic field which is applied essentially parallel to the direction of current flow. The plasma is produced by injection[14] from indium solder contacts of various cross-sectional area. The electric field strength is determined by measuring the potential difference between two very small contacts spaced a known distance apart. The series of measurements reported here involves four samples of similar bulk properties but different dimensions and contact arrangements as recorded in Table I. The waveform of the

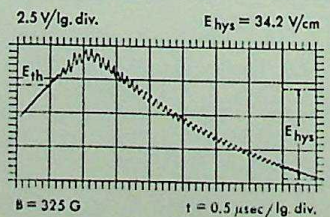


FIG. 2. An oscillogram showing hysteresis in the threshold electric field strength while the plasma is immersed in a constant magnetic field intensity.

[14] B. Ancker-Johnson, R. W. Cohen, and M. Glicksman, Phys. Rev. 124, 1745 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

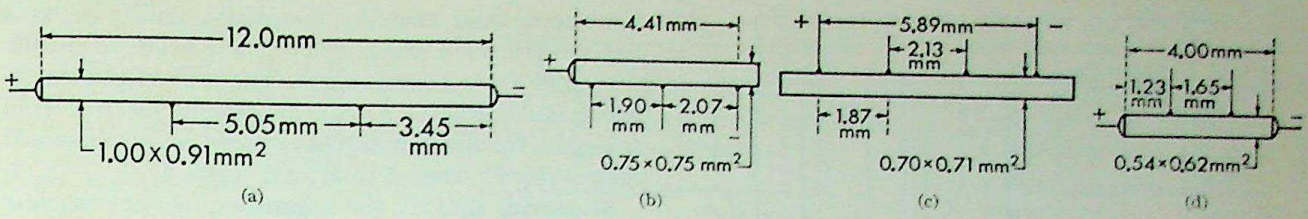


FIG. 3. Dimensions of samples. (a) 2D-2, (b) 2D2-7, (c) 2D2-8, (d) 2D2-2.

helical instability oscillations is markedly affected by the degree of parallelism (as also noted previously[5]) between the applied magnetic field and the direction of current flow, particularly in the cases of nonsymmetric current contacts (samples 2D2-7 and 2D2-8). By careful adjustment it was always possible to obtain essentially sinusoidal oscillations as illustrated by Fig. 3(d) of Ref. 11.[15]

The use of a triangular input waveform enables the simultaneous determination of the following pertinent parameters: the electric field strength threshold E_{th} at a given magnetic field strength B_{th}, and the difference in electric field strength at the onset and cessation of oscillations E_{hys}, as Fig. 2 illustrates. A separate measurement of the current-voltage characteristics (see Figs. 1 and 4 of Ref. 8 and Fig. 5 of Ref. 14, for examples) relates E_{th} to the total current and, by subtraction of the extrapolated Ohmic current, to the plasma current. Hence the boundary between stability and instability may be delineated by plotting B_{th} as a function of E_{th} with the ratio of the injected plasma density to the initial density of carriers (or background plasma density) ξ as a parameter. (See Part IV for method of calculating ξ.)

Two examples of boundary diagrams are shown in Figs. 4 and 5. The correpsonding ξ parameters for each measured point on the boundary are listed in Table I. As the electric field strength is increased with the plasma immersed in a constant magnetic field, oscillations commence at the conditions corresponding to the circles. A line connecting these points defines the boundary, a line which has no characteristic locus since the third controlling parameter ξ is arbitrarily fixed by the injection properties of the particular contacts. In practice, however, each boundary is of such a shape that the product of the threshold electric and magnetic fields is very approximately a constant. The conditions marked by the triangles on Figs. 4 and 5 indicate the observed hysteresis in electric field strength, i.e., the lower than E_{th} values at which oscillations are observed to cease.

A square-wave input was also employed whose rise and fall times (<13 nsec) are less than the period of these helical instability oscillations. With a circuit identical to that in which the triangular-waveform input is used, except for the absence of the capacitance to ground which makes the square-wave triangular, oscillations do not persist after cutoff of the square wave from an electric field strength greater than E_{th} to one less than that at which oscillations cease when a slowly falling E is applied, i.e., $E<(E_{th}-E_{hys})$. Thus, E_{hys} cannot be attributed to a slowly damping oscillation.

E_{th} is constant to within $\pm 5\%$ when the magnitude of the peak value of the applied field strength is varied from E_{th} to $2E_{th}$. The magnitude of the field strength at which the oscillations cease is constant to within $\pm 3\%$ for the same excursion into the supercritical region and, in fact, deviations can only be observed when the maximum field strengths are close to E_{th}. Variations in the time-rate-of-change of the applied field strength by as much as a factor of 4 also do not affect E_{th} and E_{hys} beyond the above-stated limits.

By completing the loops as drawn in Figs. 4 and 5 the enhancement in the magnetic induction caused by the formation of the helical instability may be deduced. The influence of the parameter ξ in deducing this enhancement is discussed in Sec. IV. Figure 4 records more typical results than Fig. 5 which shows the occurrence of some very large induction enhancements.

The larger magnitudes of the hysteresis electric field strengths, E_{hys} and larger induction increases, B_{hys},

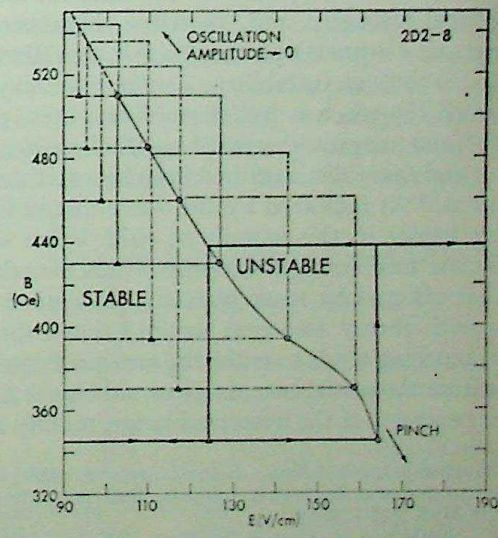


FIG. 4. A typical B–E threshold diagram showing the boundary between a quiescent and a rotating plasma. The circles which define this boundary correspond to the measured E_{th} at the onset of oscillations and the triangles indicate the measured electric field strengths at cessation, $(E_{th}-E_{hys})$. Simplified hysteresis loops corresponding to these thresholds are drawn.

[15] Small deviations (e.g., $\sim 1°$) from the alignment between the direction of current flow and the magnetic field yielding the most nearly sinusoidal oscillations do not noticeably affect E_{th} and E_{hys}.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

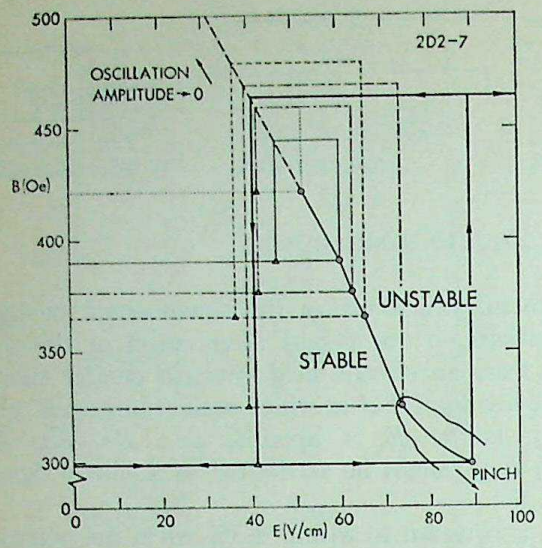


FIG. 5. A B–E threshold diagram exhibiting very large hysteresis loops.

and, therefore, the larger hysteresis loops, accompany the smaller applied magnetic fields as Figs. 4 and 5 and Table I show. The E_{hys} decreases steadily with increasing applied magnetic field but the magnitude of B_{hys} (Table I) depends on the shape of the instability boundary. If the boundary is steep at the high B end, i.e., $|dB/dE|$ is large, the B_{hys} values do not steadily decrease with increasing applied magnetic field, as in the case of samples 2D-2 and 2D2-2 whose $|dB/dE|$ at the high B end of the instability boundary are 6.5 and 7.0, respectively, compared with 3.8 and 4.2 for 2D2-8 and 2D2-7. Independent of $|dB/dE|$, however, is the fact that the hysteresis magnetic fields are smallest at the highest applied magnetic fields.

The extent of the B–E boundary over which hysteresis loops can occur is limited by two factors: At high electric field strengths and, therefore, high currents, the magnetic (Bennett) pinch effect[8,14] prevails which inhibits the helical instability. Oscillations may still occur when the pinch is well formed but these pinch-unpinch[8] and magnetothermal[16] oscillations are very different and easily distinguished from helical instability oscillations.[17] As indicated by the wavy line in Fig. 5, pinching begins in this sample at ~73 V/cm so the high electric field strength extreme of this boundary is somewhat affected by weak pinching. A detailed comparison with theory (see next section) shows that the onset of pinching tends to make the measured threshold deviate from the actual boundary toward higher E_{th}. At the other extreme of the hysteresis range, namely at the low electric field strengths, the plasma density becomes vanishingly small along with the amplitude of oscillations and the hysteresis effect. The ratio of plasma-to-background-density ξ at the lowest electric field strengths for which hysteresis can be readily observed range typically from 1 to 3% (cf. Table I).

Measuring directly the magnitudes of the magnetic fields produced by the helical paths of plasma current is difficult, principally because of the relatively small sample sizes and large currents. Inadequate electrostatic shielding plus the unavailability of a sufficiently high current source which could be appropriately modulated prevents the use of pickup coils, a highly satisfactory method for such magnetic field measurements in electron-ion plasmas.[12] Hall probes, however,

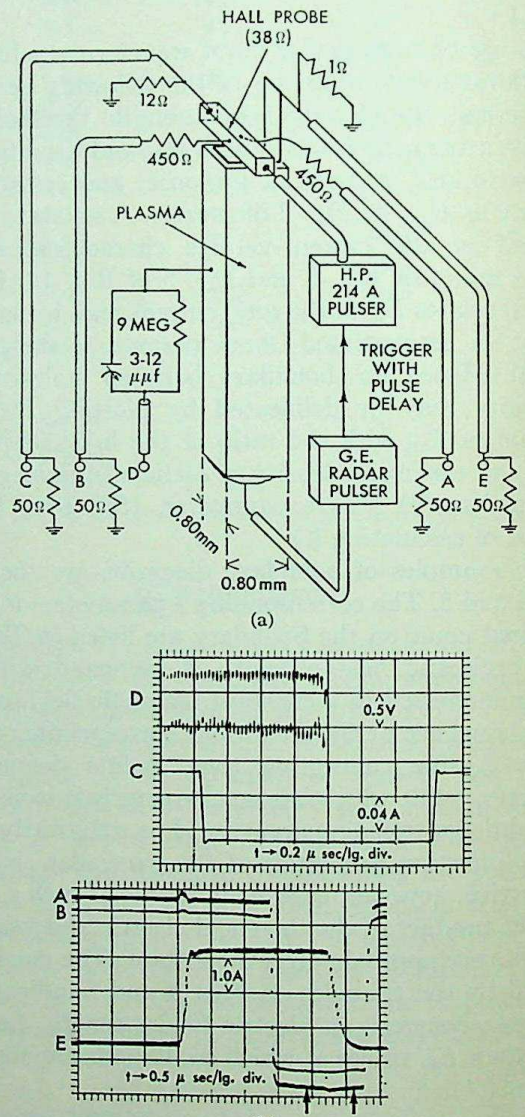


FIG. 6. (a) Experimental arrangement used to measure B_{hys} with a Hall probe. The plasma-injecting contact at the probe end is made with a gold strip. The spacing between this strip and the probe is 0.27 mm. (b) Oscilloscope traces: D shows the helical instability oscillations and their cutoff while the current in the Hall probe C remains on. The lower oscillogram has a faster sweep time and shows both the beginning and end of the current pulse in the plasma E. Traces A and B correspond to the voltages on the Hall probe side contacts.

[16] B. Ancker-Johnson, Phys. Rev. Letters 9, 485 (1962); J. E. Drummond and B. Ancker-Johnson, Bull. Am. Phys. Soc. 8, 471 (1963); 9, 318 (1964).

[17] M. C. Steele and T. Hattori [J. Phys. Soc. Japan 17, 1661 (1962)] offer an explanation of some very small amplitude oscillations observed during pinching by M. Glicksman and R. A. Powlus [Phys. Rev. 121, 1659 (1961)] in impact-ionization plasmas in n-type InSb which is related to the generation of sound waves. Impact ionization is not occurring in the present experiments.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

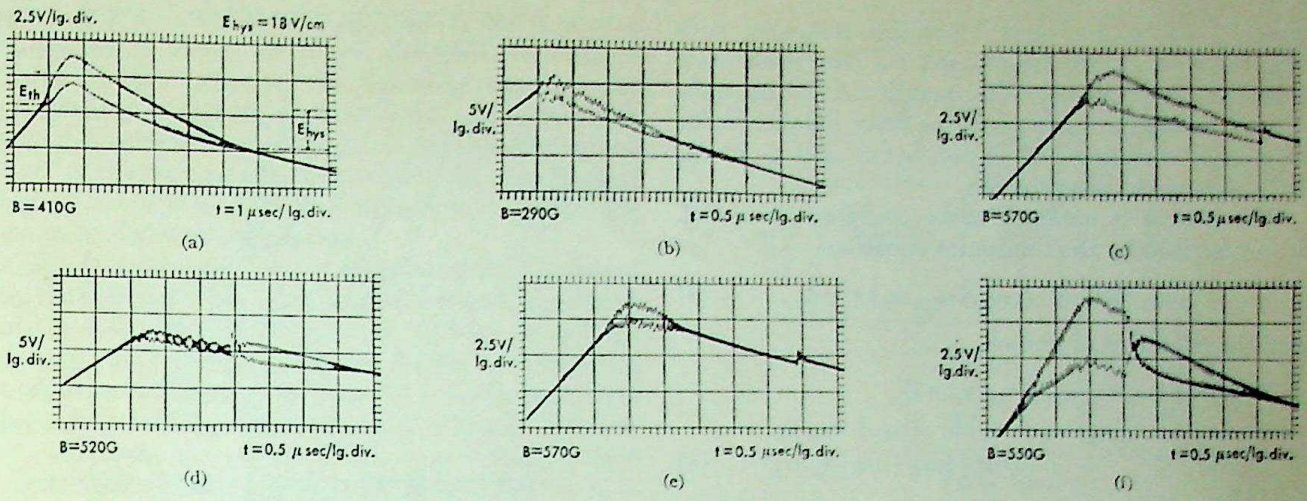


FIG. 7. Oscilloscope diagrams illustrating some of the startling properties of the oscillatory behavior.

do yield a direct estimate when the arrangement shown in Fig. 6 is employed. A small parallelepiped of n-type InSb is located orthogonally to a plasma-containing sample, so that the former senses the applied longitudinal magnetic field as well as the helix field, but not azimuthal fields produced by current flow.

The Hall probe circuit is supplied with resistances such that the transmission line is everywhere matched and the voltage contacts A and B are effectively isolated from the current path. The outputs of A and B are fed into a storage oscilloscope (Tektronics No. 564) equipped with a sampling dual-trace plug-in unit (No. 3S76). These voltages, at a time selected with the manual scan of a sampling sweep plug-in unit (No. 3T77) and indicated by the arrows at the bottom of Fig. 6(b), are taken from the output jacks of the dual-trace unit through two bucking voltages supplies and a pair of dc amplifiers. The difference between these signals is then read on a voltmeter which is directly calibrated in gauss by varying the applied longitudinal magnetic field. The bucking voltages are adjusted to maximize the sensitivity: No signal is detected when the applied magnetic field is on but no oscillations are present [right-hand arrow in Fig. 6(b)], so that only the field added by the helix is measured when the manual scan is located at the left-hand arrow. Although the sensitivity of this system is $\sim$1 G, drifting of the zero makes the measurements tedious. The results are summarized in Table II.

Certain observed aspects of the hysteresis effect are quite startling. Some are illustrated by the oscillograms of Fig. 7. The first, 7(a), shows how rapidly the amplitude of oscillations can grow to surprisingly large amplitudes in an otherwise perfectly typical hysteresis oscillogram. Oscillations can apparently cease as the electric field strength decreases to some magnitude and then with further decrease commence again, 7(b).[18] Unusually large amplitude oscillations occasionally cease suddenly with decreasing electric field strength, producing a discontinuity in the average E, 7(c). A sudden switch from moderate to larger amplitude oscillations can also occur during decreasing E, 7(d). No hysteresis or a "reverse hysteresis," an anomaly in itself, can be followed by a sudden discontinuity in the measured E at some E well below $(E_{th}-E_{hys})$, 7(e). Rare but reproducible is the oscillatory pattern displayed in 7(f). The input waveform was monitored before each of these oscillograms was made and found to be smooth as usual.

Detailed observations regarding threshold frequency (i.e., frequency at the instability boundary) cannot be made from this type of data since the frequency is a function of electric field strength in the supercritical region.[2,3,6-10]

III. THEORY

The helical instability[2,3,6,7] is a growing helical density perturbation superimposed on an unperturbed, steady-state plasma density and electric potential in a plasma column immersed in parallel (or antiparallel) electric and magnetic fields.[19] The perturbed density and potential are assumed to have the following forms, respectively,

$$n_i = n_{i,0} + f(r,z)\exp[ikz + im\theta + i\omega t], \quad (1)$$

$$U = U_0 + g(r)\exp(ikz + im\theta + i\omega t], \quad (2)$$

TABLE II. Hall probe measurements on sample 2D2-13.

Bulk properties	$B_{II}(G)$	$I_p(A)$	when $B_{helix}(G)$ at end is:	$I_p(A)$ at threshold
$p_0=3.5\times10^{14}$ cm^{-3}	340	1.9	10±2	0.4
$\mu_p=8000$ cm^2/V-sec	320	2.3	13±2	0.5

[18] A possibly related and also as yet not understood effect was noted in Ref. 8: As the magnetic field strength was increased above threshold while the total current was held constant, the oscillations became noisy and then another "mode" at a higher frequency emerged before noise ultimately resulted.

[19] F. C. Hoh and B. Lehnert, Phys. Rev. Letters 7, 75 (1961) give one simple, physical interpretation of the theory.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where f and g are small quantities compared with $n_{i,0}$, the steady-state density and U_0, the steady-state potential, respectively. The quantity k is the wave number along the z axis, m is the wave number in the azimutual direction ($m=1$ for helix) and ω is the frequency of the perturbation.

The plasma is assumed to be collision dominated. It is described by the continuity equations

$$(\partial n_{i\pm}/\partial t)\pm\nabla\cdot\{(n_{0\pm}+n_{i\pm})v_\pm\}=\gamma n_{i\pm}, \tag{3}$$

and the equations of motion

$$D_\pm\nabla n_{i\pm}=\pm(\mu_\pm/c)(n_{0\pm}+n_{i\pm})\mathbf{v}_\pm\times\mathbf{B} \pm\mu_\pm(n_{0\pm}+n_{i\pm})\mathbf{E}-(n_{0\pm}+n_{i\pm})\mathbf{v}_\pm -(\partial/\partial t)[(n_{0\pm}+n_{i\pm})\mathbf{v}_\pm]\nu_\pm^{-1}. \tag{4}$$

The last term in (4) is negligible because the collision frequencies $\nu_\pm$ are so large. The subscripts $\pm$ refer to holes and electrons, respectively; γ is the recombination and generation coefficient; n_0 is the density of initial carriers, D is the diffusion constant, μ the mobility and v the velocity. An essentially neutral injected plasma $|n_{i+}-n_{i-}|\ll|n_{i\pm}|$ with equal electron and hole temperatures is assumed. Even at relatively elevated electric field strengths ($\sim$200 V/cm) equal temperatures seems to be the most reasonable assumption.[20]

A dispersion relation for the frequency is derived from Eqs. (1) and (2) and the boundary between stability and instability occurs when $\mathrm{Im}(\omega)=0$. Using the additional condition[3,7] that the derivatives of Im (ω) with respect to k must equal 0 at the onset of instability, the electric field strength (also the wavelength and frequency of oscillation) is expressed in terms of the threshold magnetic field strength.

Holter[7] has taken into account that the flow of current causes plasma heating above the semiconductor lattice temperature, here $V_0=6.6\times10^{-3}$ eV. Hence the mobilities are expressed as

$$\mu_\pm=\mu_{0\pm}(V_0/V)^\alpha, \tag{5}$$

where $\mu_{0\pm}$ are the hole and electron mobilities, respectively, at V_0 which is less than the plasma temperature V. For InSb in the temperature range of interest $\alpha\approx0.5$. Then dimensionless threshold magnetic and electric field parameters may be expressed, respectively, as

$$y=y_0(V_0/V) \tag{6}$$

and

$$\mathcal{E}=\mathcal{E}_0(y/y_0). \tag{7}$$

With the use of these definitions

$$y_0=\mu_{0-}\mu_{0+}B_{\mathrm{th}}{}^2, \tag{8}$$

$$\mathcal{E}_0=E_{\mathrm{th}}R/V_0. \tag{9}$$

R is the radius of the plasma. Equation (7) is obtained by eliminating the temperature V when Eqs. (6) and (9) are combined. In the limit of small threshold magnetic fields, i.e., $y\ll1$, which is the case in these experiments, $\mathcal{E}$ has an asymptotic solution which may be expressed as

$$\mathcal{E}=K_0y^{-1/2}. \tag{10}$$

K_0 is dependent on the injected-plasma-to-background-density ratio ξ but not on magnetic field, hence, it may be evaluated for a particular $y\ll1$, e.g., $y_1=10^{-3}$. Then the corresponding values of $\mathcal{E}_1$ are known as a function of ξ and the plasma temperature is found by equating Eqs. (7) and (6) and using (9) and (10):

$$V=\left[\frac{y_0V_0}{y_1}\left(\frac{E_{\mathrm{th}}R}{\mathcal{E}_1}\right)^2\right]^{1/3}. \tag{11}$$

A nonlinear theory treating the hysteresis effect itself in electron-ion plasmas has been worked out by Holter and Johnson.[21] Such a theory applicable to electron-hole plasmas is as yet lacking.

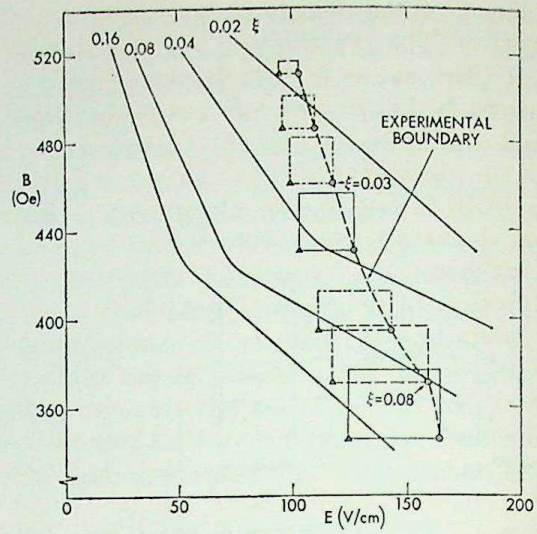


FIG. 8. B–E threshold diagrams for plasmas of constant densities in the same sample as used to obtain the data in Fig. 4. Simplified hysteresis loops corresponding to constant ξ values are shown. The experimentally determined threshold boundary (nonconstant ξ) is also indicated. The ξ values for two experimental points are identified for an illustration (see text).

IV. DISCUSSION

The parameter ξ is estimated from the readily obtainable Ohmic and total current magnitudes, respectively,

$$I_\Omega=ep_0\mu_{0+}EA, \tag{12}$$

$$I_T=I_\Omega[1+((\mu_{0-}/\mu_{0+})+1)n/p_0], \tag{13}$$

as

$$\xi\approx\frac{n}{p_0}=\frac{I_T-I_\Omega}{I_\Omega((\mu_{0-}/\mu_{0+})+1)}, \tag{14}$$

when A is the cross-sectional area of a semiconductor

[20] M. Glicksman and W. A. Hicinbothem, Jr., Phys. Rev. 129, 1572 (1963).

[21] O. Holter and R. R. Johnson, Boeing Document D1-82-0256, December 1963 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

containing an intial density of holes p_0 and a plasma of density n. The experimental ξ values are listed in Table I. (The radius of the plasma assumed for the calculations is that of the inscribed circle within each semiconductor cross section.) The calculated temperatures of the plasmas at the instability threshold ranged from 5.9×10^{-2} to 1.1×10^{-2} eV while the lattice remains at the bath temperature.

The influences of ξ on the enhancement in B caused by the presence of an helical current path are considered in the following: As E increases so does the plasma density which is injected by the current contacts, and as E exceeds E_{th} more and more of the plasma current goes into a helical path producing an enhanced B. When E decreases, the total $B=B_{applied}+B_{helix}$ remains greater than B_{th} even though the plasma current I_p is decreasing. Thus, a cessation of oscillation occurs at an $I_p<I_p$ at threshold and, therefore, at a ξ value $<\xi_{th}$. This fact and its influence on the magnitude of the deduced B_{hys} is illustrated by Figs. 8 and 9. The lines labeled with numbers represent the threshold conditions for constant ξ values. The experimental threshold points are represented by circles and the triangles refer to conditions at cessation of the oscillations, consistent with the symbols in Figs. 4 and 5. At both threshold and cessation the amplitude of the oscillations is vanishingly small (cf. Fig. 2) and, hence, the perturbation theory[7] is applicable.

Hysteresis loops are drawn in Figs. 8 and 9 corresponding to constant ξ-value boundaries. The resulting B_{hys} are smaller than those deduced for nonconstant ξ boundaries, Table I. In general, the magnitude of I_p at some operating point in the supercritical region determines the B produced by the helix. If ξ were made independent of E and constant, e.g., by supplying the plasma optically when noninjecting contacts carry the current, the smaller set of B_{hys} magnitudes would

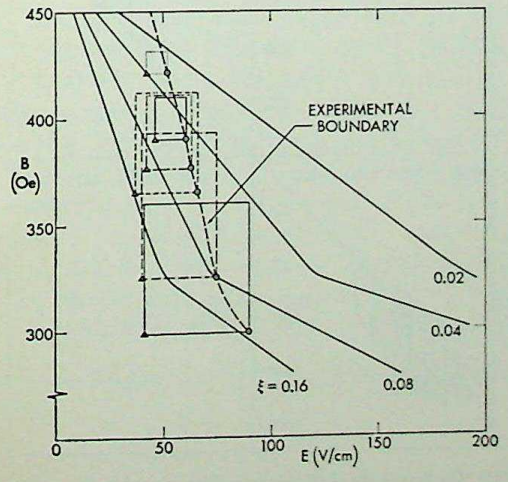


FIG. 9. B–E threshold diagrams for plasmas of constant density in the same sample as used to obtain the data in Fig. 5. Simplified hysteresis loops corresponding to ξ values are shown. The experimentally determined threshold boundary (nonconstant ξ) is also indicated.

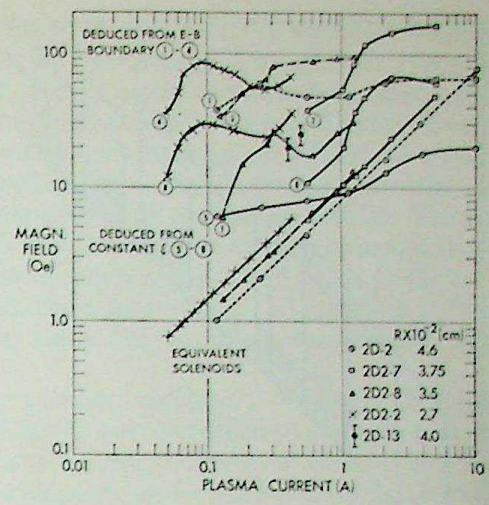


FIG. 10. Magnetic field is plotted as a function of plasma current using: (1–4) the experimentally deduced B_{hys} values (nonconstant ξ); (5–8) B_{hys} values derived from constant ξ conditions; and calculated B values produced by solenoids of equivalent turns ratio. The two points with error brackets represent the results of Hall probe measurements.

result. Conversely, the steeper the I–V characteristic and, therefore, the more rapid the increase of plasma density with E, the steeper the B–E boundary and the larger resulting B_{hys}. Sample 2D2-7 compared with 2D2-8 is an example of the latter case (Figs. 5 and 4).

Reference to a particular set of operating conditions (see Figs. 4 and 8) clarifies these statements: A threshold for the helical instability in sample 2D2-8 is at $B=370$ G, $E=159$ V/cm and $\xi=0.08$. The plasma becomes quiescent again at $E=E_{th}-E_{hys}=117$ V/cm and: (a) for a constant plasma density and therefore $\xi=0.08$ at $B=B_{th}+B_{hys}=370+26=396$ G; (b) for a plasma density dictated by the properties of the current injecting contacts $\xi=0.03$ at $B=370+91=461$ G.

The magnitude of the fields produced by the helices are plotted in Fig. 10 as a function of the plasma current at the threshold. Curves 1 through 4 refer to the experimentally determined threshold boundaries and 5 to 8 to boundaries of constant plasma density. The shape of each of these curves is controlled, as stated earlier in connection with Figs. 4 and 5, by the arbitrary properties of the injection contacts. The two points with error brackets represent the results of the Hall probe measurements. These two points are plotted at the threshold plasma current, consistent with the rest of the data points, although the measurements were made at higher currents, Table II. Also graphed in this figure are the magnetic fields that solenoids would produce which possess the same radius and number of turns per cm as the helical instability in the four samples. The turns ratios of the equivalent solenoids are derived from Holter's calculation of the wavelength of the helix at threshold (Fig. 1, Ref. 7). The dimensionless wavelength λ/R is nearly constant for all the present measurements; it varies only between 3.2 and 3.6. The radius R (Table I) has a range of less than a factor of 2, hence

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

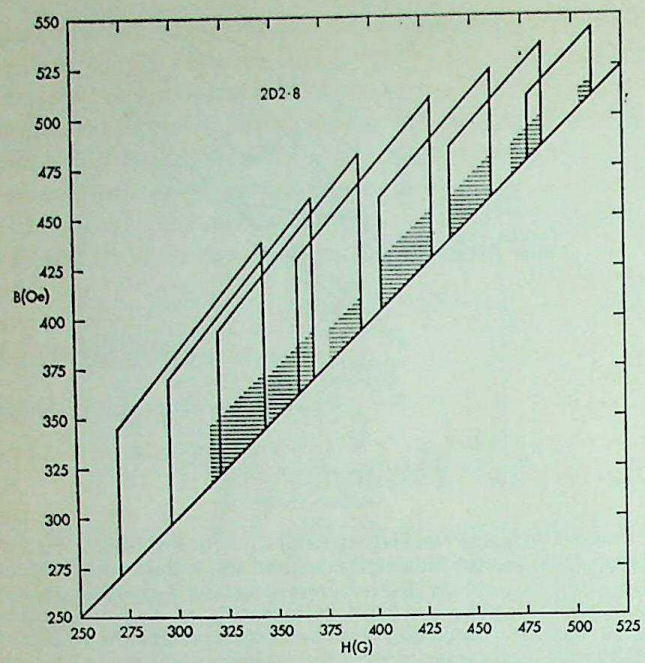


FIG. 11. The first quadrant of deduced B–H curves. The larger loops correspond to Fig. 4 and the smaller (hatched) to Fig. 8.

the fields produced by the equivalent solenoids are essentially determined by the magnitude of the current they carry. The equivalent solenoid results in Fig. 10 are obtained with the assumption that all the plasma current goes into the helix at threshold which, of course, is not compatible with the perturbation theory.

Figure 10 indicates that the equivalent solenoids produce far less B than their counterparts in the plasma at low I_p even though it was assumed that they carry all the plasma current. The fields deduced from the boundary measurements, however, result from excursions into the supercritical region, hence curves 1 to 8 should be shifted toward higher currents for comparison with the solenoid curves. The fields sensed by the Hall probes, when plotted against the I_p they carried during the measurements (Table II), agree well with the fields produced by equivalent solenoids conducting the same current. The agreement may be so good fortuitously since the λ of the helix in the supercritical region may not be constant in the range of these threshold magnetic fields. There is experimental evidence to suggest it is not constant[8,22] but a nonlinear theory for electron-hole plasmas, as stated earlier, is not available.

The Hall probe measurements were made on a sample with radius somwhat larger than 2D2–7. At $I_p = 0.5$ A the B_{hys} deduced from the experimentally determined threshold boundary is $\sim$35 G whereas that deduced from the constant-plasma-density boundary is $\sim$10 G. Since the Hall-probe value of $\sim$26 G is likely to be too small because the measurements were made a finite distance from the end of the plasma helix [Fig. 6(a)], agreement with the experimentally determined B_{hys} value is indicated.

The magnitude of B_{hys} is very large compared with the part of it associated with the diamagnetism normally exhibited by a plasma. Recently Moore and Kessler[23] have measured the magnetic moment of an optically injected plasma in Ge. For a plasma density equal to 10^{15} cm^{-3} and an essentially infinite surface recombination velocity, a condition corresponding to the highest magnetic moment, they find a diamagnetic moment density of $\sim 2\times 10^{-5}$ cgs units at 1 kG. The plasmas utilized in the present experiments can be expected to exhibit an even smaller moment since the plasma densities are 1 to $2\frac{1}{2}$ orders of magnitude smaller and the surface recombination velocity is finite. Hence essentially all of the deduced B_{hys} is associated with the paramagnetism produced by the helical instability. The first quadrant of the resulting B–H curves corresponding to Figs. 4 and 8 (the more typical set of hysteresis loops) is shown in Fig. 11. The larger loops correspond to the experimentally determined threshold boundary and the smaller loops (hatched) to the constant ξ boundary. The full B–H diagrams for sample 2D2-7, which exhibited the largest loops, are shown in Fig. 12. The loops are again drawn with straight sides for simplicity but their shapes cannot be ascertained by the present experiments.

The sign of the enhanced induction associated with the helical instability is positive as stated in the introduction, according to Johnson's measurements in electron-ion plasmas[12] and as verified by the described Hall probe measurements. This switch from diamagnetism to paramagnetism is completely compatible with

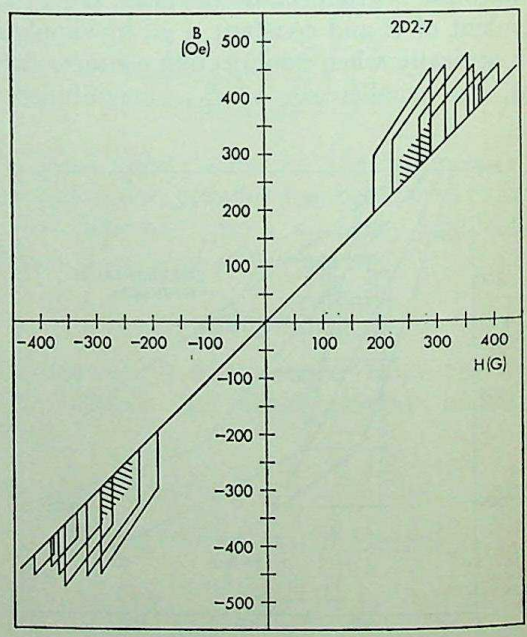


FIG. 12. Deduced B–H diagrams. The larger loops correspond to Fig. 5 and the smaller (hatched) to Fig. 9.

[22] The only published wavelength measurements known to the author are in Ref. 9. These results yield a λ/R 4 to 5 times Holter's calculated values. This makes the equivalent solenoid fields even smaller.

[23] A. R. Moore and J. O. Kessler, Phys. Rev. 132, 1494 (1963). The author is grateful to Dr. Moore for discussing their results prior to publication.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

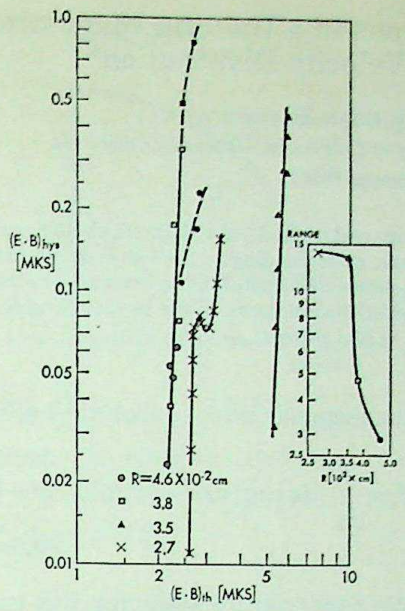


FIG. 13. The hysteresis loop energy as a function of the input energy at threshold for the experimentally determined boundaries, i.e., ξ nonconstant. The solid points represent hysteresis energy in a weakly pinched plasma. The insert shows the range of hysteresis loop energy as a function of plasma radius.

the helical instability theory[9b] which states that the the helical instability theory which states that the sense of the helix is positive (negative) with respect to the applied magnetic field when the applied E is parallel (antiparallel), and is independent of the drift direction of the helical density perturbation. Enhancement of the magnetic field results and, therefore, paramagnetism. That paramagnetism results is also evident from simple physical reasoning: Since the application of a B, large compared to B_{th}, causes a shift in the equilibrium position of the plasma toward its container walls, the magnetic pressure inside the helix must be greater than that outside, thus the magnetic field produced by the helix must *add* to the applied B.[24]

The effect of the plasma radius on the hysteresis is illustrated by Fig. 13 in which a measurement of the "hysteresis loop energy" is plotted against the "input energy" required to achieve oscillations for the four different plasma radii. The input energy necessary to produce hysteresis is quite constant as expected from both theory and the experimental results shown in Figs. 4 and 5. The theoretical curves of E_{th} as a function of B_{th}^2 have a slope of -0.5, hence for any one value of ξ the product $E_{th}\times B_{th}$ is a constant as illustrated by the hyperbolic boundaries for constant ξ values drawn in Figs. 8 and 9. Since ξ varies in the experiments (Table I), this product is only approximately constant. Of considerable practical interest is the fact that very little change in input energy is necessary for very large changes in loop energy. The greatest range in loop energy, defined as the ratio of the loop energy at the lowest $(E\times B)$ at which observable hysteresis occurs to the loop energy at the highest $(E\times B)$ at which noise-free oscillations occur, is produced by the smallest diameter plasma, if pinching is avoided. The insert in Fig. 13 indicates that the range saturates with decreasing plasma radius at $\sim 3\times 10^{-2}$ cm. If weak pinching is allowed, there appears to be an optimum intermediate plasma radius for obtaining the greatest hysteresis range.

The startling oscillatory behavior illustrated by Figs. 7(b)–(f) may be related to a complicated relationship between the direction of current flow and the applied magnetic field direction, since such behavior has not been observed in the two samples investigated which possess symmetrically located plasma injectors, namely samples 2D-2 and 2D2-2.

V. CONCLUSION

The hysteresis in electric field strength E accompanying the onset and cessation of the helical instability can exceed 45 V/cm and 50% of the applied E at threshold. The attendant hysteresis in magnetic field intensity B, deduced from measurements of the boundary conditions between a quiescent and a rotating plasma which agree well with theory, can exceed 160 G and 55% of the applied B at threshold. Hall probe measurements of the B produced by the helical path of plasma current tend to verify the deduced B_{hys} values both in magnitude and sign (paramagnetic). The resulting, novel B–H curves consist of loops which are displaced from the origin. An infinite number of hysteresis loops can be generated by a small variation in input energy $(E_{th}\times B_{th})$. The extent in E (or B) over which loops can occur is limited at the high E (low B) end by the onset of current pinching and at the low E (high B) end by vanishing plasma current and, hence, vanishing hysteresis. The existence of the largest loops at low magnetic fields is in agreement with Holter and Johnson's nonlinear theory[21] of the helical instability in electron-ion plasmas.

These loops are easily and reproducibly obtained in p-type InSb at 77°K.

ACKNOWLEDGMENTS

The author is grateful to Dr. James E. Drummond for very helpful conversations and Robert W. Boice and Michael F. Berg for invaluable technical assistance.

[24] The author is grateful to H. P. Furth for suggesting this line of reasoning.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Stability of Longitudinal Oscillations in a Uniform Magnetized Plasma with Anisotropic Velocity Distribution*

LAURENCE S. HALL AND WARREN HECKROTTE
Lawrence Radiation Laboratory, University of California, Livermore, California

(Received 13 January 1964)

We consider the longitudinal oscillations of a uniform magnetized plasma in which the particles are allowed to have anisotropic distributions in velocity space, corresponding to the two temperatures $T_\perp$ and $T_\parallel$ perpendicular and parallel to the magnetic field. It is shown that no instabilities can exist unless $T_\perp > T_\parallel$ for at least one of the species of plasma particles. An additional theorem, of use in surveying for possible unstable roots to the dispersion equation when $T_\perp > T_\parallel$, is also proved.

THE stability of longitudinal oscillations of a uniform plasma with anisotropic velocity distribution in the presence of a uniform magnetic field has been studied by a number of authors[1–3] in various approximations. We consider here a multicomponent nonrelativistic plasma in which the zeroth-order distribution function may be written as a bi-Maxwellian, viz.,

$$f_{0j}=(2\pi\kappa T_{\perp j}/m_j)^{-1}(2\pi\kappa T_{\parallel j}/m_j)^{-1/2} \times\exp\left[-\frac{m_j v_\perp^2}{2\kappa T_{\perp j}}-\frac{m_j v_\parallel^2}{2\kappa T_{\parallel j}}\right], \quad (1)$$

and show that no instability exists unless $T_{\perp j}>T_{\parallel j}$ for at least one species. (Here $\parallel$ and $\perp$, respectively, denote directions parallel and perpendicular to the magnetic field B_0; j denotes the species of particle; and m, T, and κ are, respectively, mass, temperature, and Boltzmann's constant.) An additional theorem of use in locating possible instabilities when $T_{\perp j}>T_{\parallel j}$ is also proved.

Assuming that all perturbations vary in space and time according to the form $\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$, where $\mathbf{k}$ is real and ω complex, and $\mathrm{Im}(\omega)>0$, one can follow step by step the prescription of Bernstein[4] in obtaining the dispersion relation for longitudinal oscillations. After a straightforward but laborious computation, one may write

$$1+\textstyle\sum_j \omega_{pj}^2\Omega_j^{-2}F_j=0, \quad (2)$$

where $\omega_{pj}=(4\pi Z_j^2e^2n_j/m_j)^{1/2}$=plasma frequency, $\Omega_j=|Z_je|B_0/m_jc$=cyclotron frequency, and the sum is over particle species. If we write $\omega_j=\omega/\Omega_j$; $\gamma_j=k(2\kappa T_{\perp j}/m_j\Omega_j^2)^{1/2}=k\cdot$(radius of gyration); $t_j^2=T_{\parallel j}/T_{\perp j}$; and take $n_\parallel$ and $n_\perp$, respectively, to be the direction cosines of the propagation vector parallel and perpendicular to the magnetic field, so that $n_\parallel^2+n_\perp^2=1$, we have

$$F_j=\int_0^\infty dx\,\exp\{i\omega_j x-\tfrac{1}{4}\gamma_j^2t_j^2n_\parallel^2x^2-\tfrac{1}{2}\gamma_j^2n_\perp^2(1-\cos x)\} \times[n_\parallel^2x+n_\perp^2\sin x]. \quad (3)$$

We now drop the subscript j as understood, and consider the integral in (3) noting that it may be rewritten after an integration by parts of the first term to read

$$F=2\gamma^{-2}t^{-2}+2\gamma^{-2}t^{-2}F_A+(1-t^{-2})n_\perp^2F_B, \quad (4)$$

where

$$F_A=i\omega\int_0^\infty dx\,\exp\{i\omega x-\mu x^2-\lambda(1-\cos x)\}, \quad F_B=\int_0^\infty dx\,\exp\{i\omega x-\mu x^2-\lambda(1-\cos x)\}\sin x, \quad (5)$$

and we have set $\mu=\frac{1}{4}\gamma^2t^2n_\parallel^2$ and $\lambda=\frac{1}{2}\gamma^2n_\perp^2$. All of the parameters are real except ω which is complex and which, for an instability, has positive imaginary part. We propose to show that, when $\mathrm{Im}(\omega)>0$,

$$\mathrm{Re}(\omega)\cdot\mathrm{Im}(F_A)\geq 0, \quad \mathrm{Re}(\omega)\cdot\mathrm{Im}(F_B)\geq 0. \quad (6)$$

Since both $\mathrm{Im}(F_A)$ and $\mathrm{Im}(F_B)$ are antisymmetric in $\mathrm{Re}(\omega)$, it is only necessary to prove (6) for ω in the first quadrant of its complex plane. We then use the fact that if $f(\omega)$ is an analytic function of ω in the region of consideration, the map of any closed contour C in the ω plane onto the f plane encircles all values of f attained by letting ω assume values within C. We choose C as in Fig. 1, and note that F_A and F_B both vanish on the infinite quarter circle. Further, for ω on the imaginary axis both $\mathrm{Im}(F_A)$ and $\mathrm{Im}(F_B)$ vanish. Thus, we need only show that $\mathrm{Im}(F_A)$ and $\mathrm{Im}(F_B)$ are sign-definite for ω real and positive in order to show they are sign-definite for ω anywhere in the first quadrant.

First, consider $\mathrm{Im}(F_A)$ and use the identity

$$e^{\lambda\cos x}=\sum_{l=-\infty}^{\infty} I_l(\lambda)e^{ilx}, \quad (7)$$

where the $I_l(\lambda)$ are Bessel functions of the first kind

* This work performed under the auspices of the U. S. Atomic Energy Commission.

[1] E. G. Harris, J. Nucl. Energy **C2**, 138 (1961), and Oak Ridge National Laboratory Report ORNL-2728, 1959 (unpublished).
[2] Y. Ozawa, I. Kaji, and M. Kito, J. Nucl. Energy **C4**, 271 (1962), and references cited therein.
[3] T. Kammash and W. Heckrotte, Phys. Rev. **131**, 2129 (1963); **132**, (1963).
[4] I. Bernstein, Phys. Rev. **109**, 10 (1958), Appendix II.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

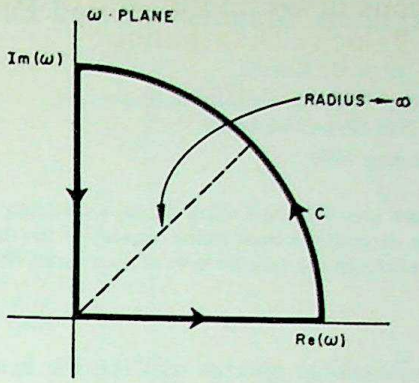


FIG. 1. Contour for the examination of various integrals.

and imaginary argument, obtaining for $\omega>0$

$$\mathrm{Im}(F_A)=\omega(\pi/4\mu)^{1/2}\sum_{l=-\infty}^{\infty}e^{-\lambda}I_l(\lambda)\times\exp\{-(\omega+l)^2/4\mu\}>0. \quad (8)$$

Similarly, with the use of (7) together with the fundamental relation $I_{l-1}(\lambda)-I_{l+1}(\lambda)=2lI_l(\lambda)/\lambda$, one can also write that for $\omega>0$

$$\mathrm{Im}(F_B)=\lambda^{-1}(\pi/\mu)^{1/2}\sum_{l=1}^{\infty}e^{-\lambda}I_l(\lambda)\times\exp\{-(\omega^2+l^2)/4\mu\}l\sinh(l\omega/2\mu)>0. \quad (9)$$

According to our previous discussion, (8) and (9) establish (6). Now consider the dispersion relation (2) which can be satisfied only if either (a) $\mathrm{Im}(F)$ vanishes for all species or (b) $\mathrm{Re}(\omega)\cdot\mathrm{Im}(F)<0$ for at least one species.[5] Condition (a) can be satisfied in general only when $\mathrm{Re}(\omega)=0$, which, if unstable, corresponds to purely growing waves. However, as pointed out by Kammash and Heckrotte,[3] $\mathrm{Re}(F)$ is positive definite in this limit and no purely growing modes can exist. Hence, an instability can only occur if $\mathrm{Re}(\omega)\cdot\mathrm{Im}(F)<0$ for at least one species. From (4) and (6), then, we have the result that no instability exists unless $t^2=T_{\parallel}/T_{\perp}<1$ for at least one species in the plasma.

It is also of interest to be able to determine whether or not roots to the dispersion relation occur when $t^2<1$, and if they exist, where they lie. In this connection, the labor involved in exploring the ω plane can often be reduced by utilization of the theorem, which we refer to as (10):

If $Re(\omega)\cdot Im[F(\omega)]$ is non-negative for ω real and lying in the interval $l\leq\omega\leq l+1$, where l is an integer, it is also non-negative in the entire strip $l\leq Re(\omega)\leq l+1$, $Im(\omega)\geq 0$. (10)

Hence, if one can show that $\mathrm{Re}(\omega)\cdot\mathrm{Im}(F)\geq 0$ for some integral range of ω on the real axis, one also can eliminate the possibility of unstable roots appearing anywhere in the strip above that range.

In order to prove (10), we rewrite (3) in the form[6]

$$F=n_{\parallel}^2F_C+n_{\perp}^2F_B, \quad (11)$$

where

$$F_C=\int_0^{\infty}xdx\exp\{i\omega x-\mu x^2-\lambda(1-\cos x)\}, \quad (12)$$

and investigate F along the contours $\mathrm{Re}(\omega)=l$. We have already proved $\mathrm{Re}(\omega)\cdot\mathrm{Im}(F_B)>0$, and so we need only prove $l\cdot\mathrm{Im}(F_C)>0$ when $\mathrm{Re}(\omega)=l$ to establish (10). Setting $\omega=l+i\sigma$ and again using the expansion (7), we can rewrite (12) after a little manipulation to read

$$\mathrm{Im}(F_C)=\sum_{k=1}^{\infty}e^{-\lambda}[I_{k-l}(\lambda)-I_{k+l}(\lambda)]\times\int_0^{\infty}xdxe^{-\sigma x-\mu x^2}\sin kx. \quad (13)$$

Consider the integral $K(\omega')=\int_0^{\infty}xdx\exp\{i\omega' x-\mu x^2\}$, $\mathrm{Im}(\omega')>0$, whose imaginary part is antisymmetric in $\mathrm{Re}(\omega')$. Evaluating K along the contour C in Fig. 1, again we find $\mathrm{Im}(K)=0$ along the imaginary axis and the infinite quarter circle. Along the real axis $k\cdot\mathrm{Im}[K(k)]=\frac{1}{4}\pi^{1/2}k^2\mu^{-3/2}\exp(-k^2/4\mu)\geq 0$. Hence k times the integral in (13), which is just $\mathrm{Im}[K(k+i\sigma)]$, is non-negative. Moreover, the Bessel function $I_k(\lambda)$ is a monotonically decreasing function of the absolute value of its order, so that $[I_{k-l}(\lambda)-I_{k+l}(\lambda)]\geq 0$. Hence $\mathrm{Re}(\omega)\cdot\mathrm{Im}(F_C)>0$ and (10) is proved.

In conclusion, we have shown that in a uniform magnetized plasma whose particles are distributed anisotropically in velocity space in zero order according to the bi-Maxwellian form (1), no unstable longitudinal oscillations exist unless $T_{\perp}>T_{\parallel}$ for at least one species. In addition, instabilities can exist when $T_{\perp}>T_{\parallel}$ only for complex frequencies lying in the strip $\mathrm{Im}(\omega)>0$, $l<\mathrm{Re}(\omega)<l+1$, where l is an integer, and where $\omega\,\mathrm{Im}(F)<0$ for some real ω lying in the range $l<\omega<l+1$.

[5] We consider that the plasma has at least two species of particles. If one species only is principally responsible for the motion, we always consider here that the small but vanishing contribution from the neglected particles is in the stabilizing sense, as will almost always correspond to the physical situation. This means that any instability arising as a limiting case of (b) must be such that $\mathrm{Im}(F)\rightarrow 0^-$ for the active species.

[6] We may note in passing that since F may be written in either of the forms (4) and (11), the inequalities (6) show that there can be no instabilities whenever either $n_{\parallel}=0$ or $n_{\perp}=0$. This is an extension of the result noted by Harris (Ref. 1) for the special case $T_{\parallel}=0$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Lattice Vibrational Spectra of Beryllium, Magnesium, and Zinc

JAMES A. YOUNG AND JUAN U. KOPPEL
General Atomic Division of General Dynamics, John Jay Hopkins Laboratory for Pure and Applied Science, San Diego, California

(Received 27 January 1964)

The lattice vibrational spectra of Be, Mg, and Zn have been obtained using the root sampling method. The force constants used in the calculations were those determined from measurements of the dispersion relations by slow neutron-scattering techniques. Calculations of the specific heat of each metal show good agreement with measured values.

I. INTRODUCTION

IN recent years the scattering of slow neutrons in solids has proved valuable as a tool for investigating the dispersion relations of lattice waves in crystals. In particular, measurements have been made on beryllium, magnesium, and zinc, all having the hexagonal close-packed lattice structure. In the first case (Be), Schmunk *et al.*[1] have determined the force constants for a central force model with interactions out to fifth nearest neighbors. For Mg, Collins[2] determined the force constants for both a tensor force model and a central force model with interactions out to fourth nearest neighbors. A similar experiment by Borgonovi *et al.*[3] has given the force constants of zinc for the tensor force model used for Mg.[2] Using these experimentally determined force constants, we have computed the lattice vibrational spectra and specific heat for each of the above metals.

II. RESULTS

The vibrational spectrum of a crystal $\rho(\omega)$ is defined such that $\rho(\omega)d\omega$ is the fraction of frequencies ω between ω and $\omega+d\omega$. The method used here to approximate $\rho(\omega)$ is to sample the frequencies $\omega(\mathbf{k})$ for a large number of values of the wave vector $\mathbf{k}$ in the first Brillouin zone of the crystal. The frequencies obtained must be weighted according to their degeneracy in the zone, and a histogram constructed of the number of frequencies falling in a given frequency interval.

In the Born-von Kármán[4] theory of lattice dynamics the frequencies ω are determined by the following secular equation:

$$|D(\mathbf{k})-M\omega^2 I|=0, \tag{1}$$

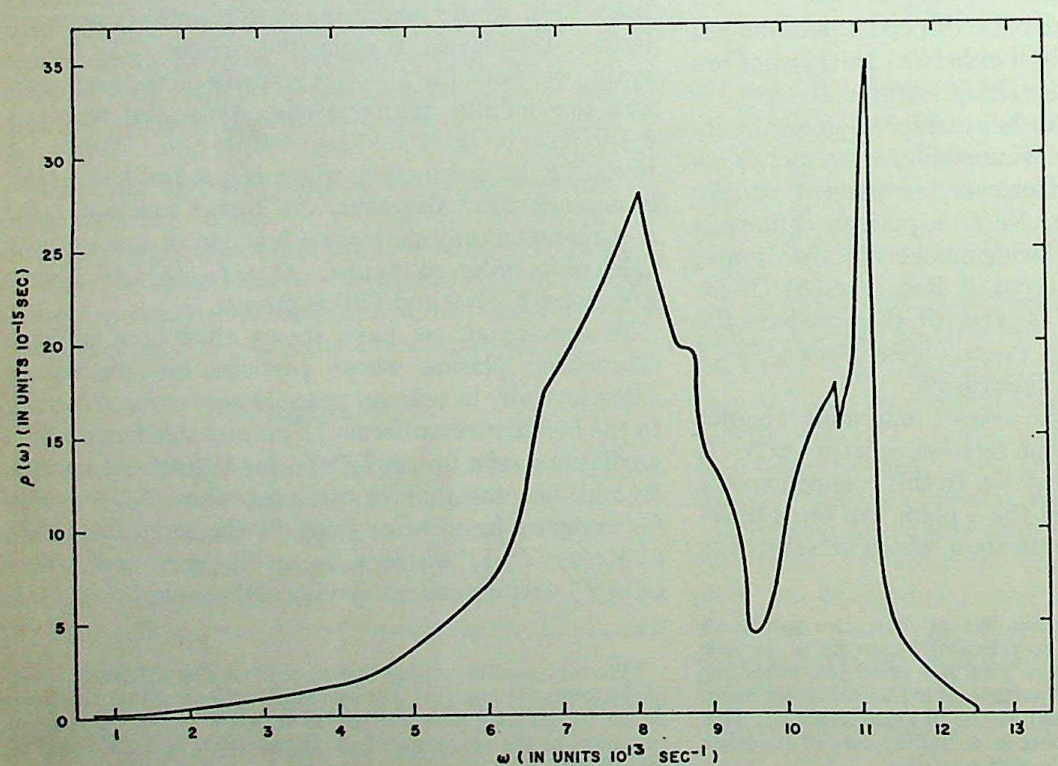


FIG. 1. Lattice vibrational spectrum of beryllium for a sampling of 10 261 points in the reduced Brillouin zone.

[1] R. E. Schmunk, R. M. Brugger, P. D. Randolph, and K. A. Strong, Phys. Rev. 128, 562 (1962).
[2] M. F. Collins, Proc. Phys. Soc. (London) 80, 362 (1962).
[3] G. Borgonovi, G. Caglioti, and J. J. Antal, Phys. Rev. 132, 683 (1963).
[4] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, London, 1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

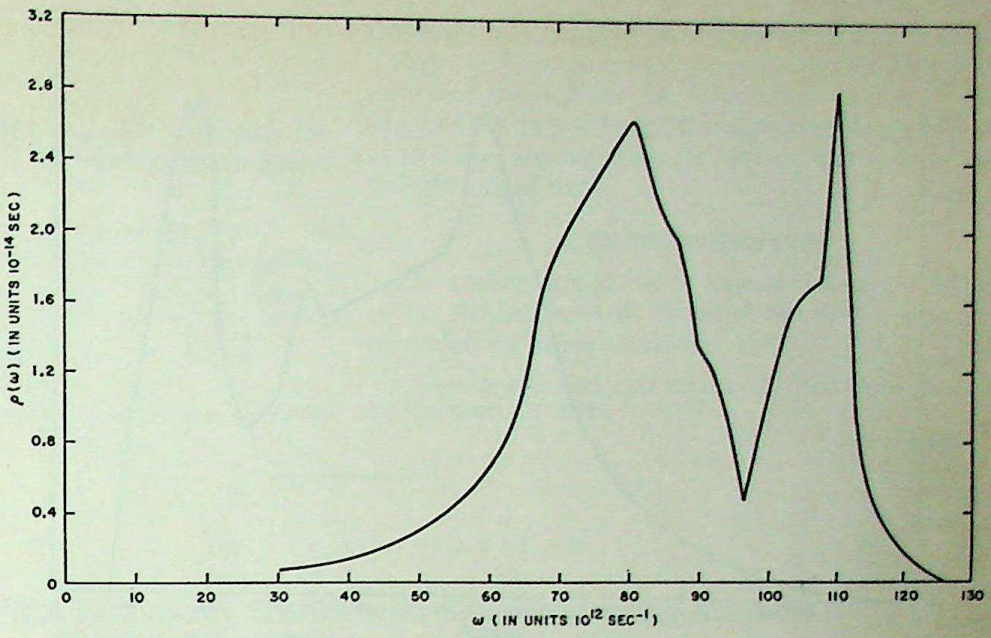


FIG. 2. Lattice vibrational spectrum of beryllium for a sampling of 2413 points in the reduced Brillouin zone.

where M is the atomic mass and I the unit matrix. The matrix $D(\mathbf{k})$ is 6×6 for hexagonal close-packed crystals and depends upon the force constants of the lattice. Equation (1) has been solved for equally spaced points in that region which is one twenty-fourth of the first Brillouin zone, and irreducible under the symmetry operations that leave the roots of the secular equation invariant. The points are then weighted as to whether or not they are on the faces, edges, etc., of this reduced zone. For beryllium, the calculation was performed for 10 261 equally spaced points ($\mathbf{k}$ values) in the reduced zone, and the spectrum shown in Fig. 1. Figure 2 shows, for comparison, the same calculation for 2413 equally spaced points. We see that greatly increasing the number of points does not significantly change the vibrational spectrum.

In the case of magnesium, the calculation was performed for a sampling of 2413 equally spaced points in the reduced zone, for both the central and tensor force models. Figures 3 and 4 show the vibrational spectra for central and tensor forces, respectively.

The vibrational spectrum of zinc was also obtained

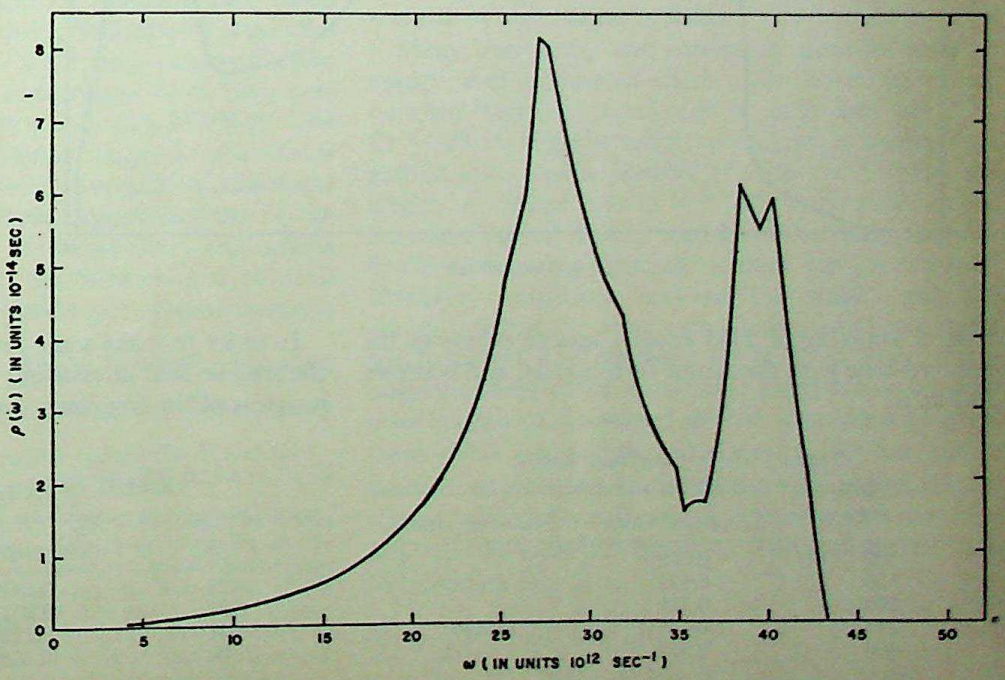


FIG. 3. Lattice vibrational spectrum of magnesium using the central force model.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

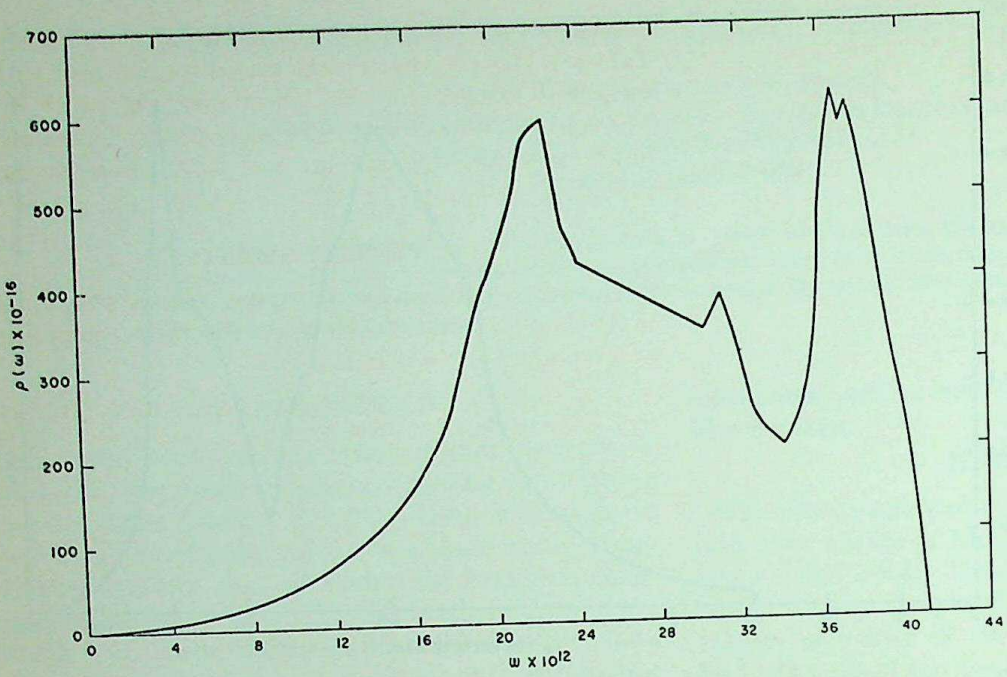


FIG. 4. Lattice vibrational spectrum for magnesium using the tensor force model.

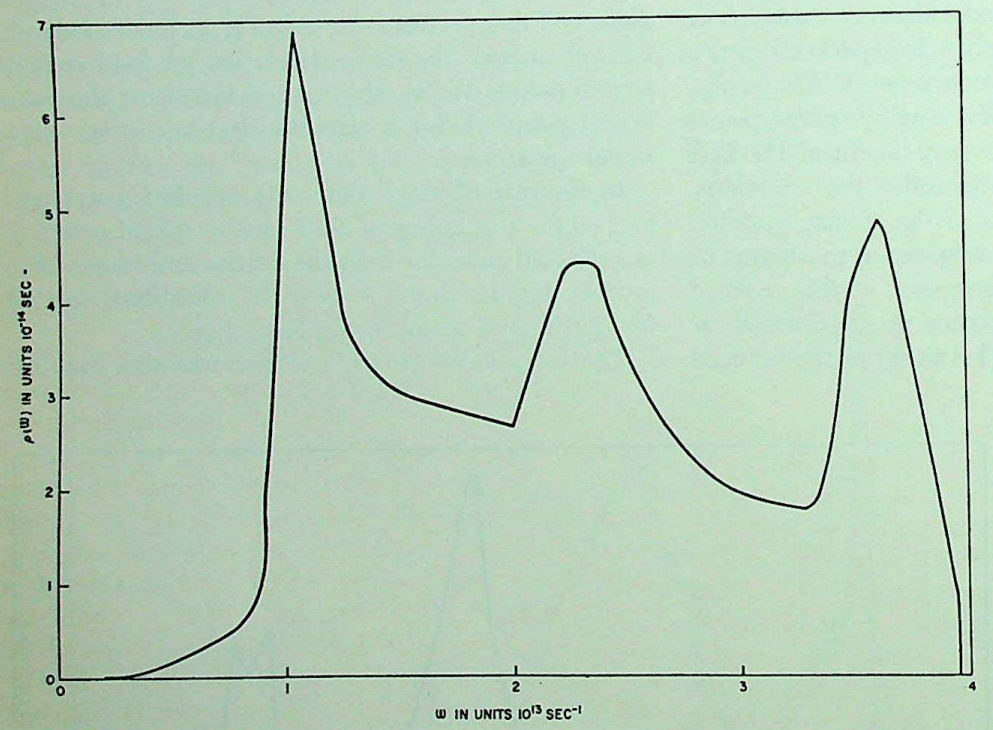


FIG. 5. Lattice vibrational spectrum of zinc using a tensor force model.

from a sampling of 2413 equally spaced points in the reduced zone with the tensor force model, and is shown in Fig. 5.

In order to make some comparison with experiment, the specific heat at constant volume C_V, which is given in terms of the frequency spectrum by

TABLE I. C_V for beryllium (cal/g).

T (°K)	Computed value	Measured value
100	0.070	
150	0.195	
200	0.311	
293	0.455	0.466
400	0.535	

TABLE II. C_V for magnesium (cal/g).

T (°K)	Computed value	Measured value
123	0.1839	0.1767
153	0.2023	0.2025
223	0.2236	0.2228
300	0.2329	

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$$\frac{C_V}{3R}=\int_0^{\omega_{\max}} d\omega\left(\frac{\omega}{T}\right)^2\frac{e^{\omega/T}}{(e^{\omega/T}-1)^2}\rho(\omega), \qquad (2)$$

where R is the gas constant, $\omega_{\max}$ the maximum frequency of the crystal, and T the temperature measured in energy units, has been computed for all three metals. Tables I–III compare the calculated specific heat with measured values,[5] for the vibrational spectra of Be, Mg, and Zn, respectively. The agreement is quite good in all cases, although specific heat comparisons do not provide a good test.

TABLE III. C_V for zinc (cal/g).

T (°K)	Computed value	Measured value
72	0.0538	0.0573
173	0.0826	0.0814
300	0.0882	

ACKNOWLEDGMENTS

The authors would like to express their appreciation to D. Naliboff and D. Houston for their competent programming of the equations.

[5] *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio, 1963).

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Electron Spin Resonance Studies with Superconducting Magnets: The Spectrum of Dy^{3+} and Sm^{3+} in CaF_2

W. Low*

Department of Physics† and National Magnet Laboratory,‡ Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 28 January 1964)

An electron spin resonance spectrometer has been constructed using strong magnetic fields up to 46 000 G generated by a superconducting magnet. It is being used for the measurement of the magnetic properties of transition elements in ionic crystals such as (a) small g factors, (b) energy level separations up to 20 cm^{-1}, (c) isotropic and anisotropic exchange terms in spectra of pairs of paramagnetic ions, (d) spin-lattice relaxation times at high magnetic fields. The method is illustrated in the measurements of (1) the energy level separation of $\Gamma_8-\Gamma_7$ levels of Dy^{3+} in CaF_2 which is found to be $7.3_5\pm0.2$ cm^{-1}; (2) the g factors of the tetragonal spectrum of Sm^{3+} in CaF_2 which are found as $g_\parallel=0.907\pm0.01$, $g_\perp=0.544\pm0.015$.

INTRODUCTION

ELECTRON spin resonance methods have been widely used in determining the magnetic properties of transition elements in ionic and semiconducting crystals.[1] Most of these experiments have been performed with magnetic fields of less than 20 000 G.[2] The availability of superconducting magnets with fields of the order of 50 000 G permit a number of additional experiments of interest. These experiments are of the kind: (a) Measurements of energy level separations of the order of 3–20 cm^{-1}. (b) Measurement of small g factors, i.e., $g\ll1$. (c) Measurement of some isotropic or anisotropic exchange splittings of neighboring paramagnetic ions. (d) Spin-lattice relaxation phenomena in high magnetic fields.

Many iron group and rare-earth elements have an energy level scheme in which a few levels are closely bunched together. Examples of such ions are V^{3+}, Cr^{4+} (d^2) in trigonal fields which have a singlet as a ground state and a doublet at 7–10 cm^{-1} above the singlet. A similar case is Fe^{2+} (d^6) in noncubic fields. In some crystal hosts, transitions between the two levels of the non-Kramers doublet can be induced. However, transitions between the singlet and the doublet cannot be seen with conventional resonance spectrometers and magnetic fields. With very strong magnetic fields one of the doublet levels can be brought close enough to the singlet so that transitions between these levels can be effected. One usually can find a suitable magnetic-field orientation with respect to the crystal axis which will permit transitions of the type $\Delta S_z=\pm1$ of sufficient intensity. The appropriate spin Hamiltonian is

$$\mathcal{H}=g_{\parallel}\beta HS_z+g_{\perp}\beta(H_xS_x+H_yS_y)+D(S_z^2-\tfrac{1}{3}S(S+1))+E(S_x^2-S_y^2), \qquad (1)$$

* Permanent address: The Hebrew University, Jerusalem, Israel.

† Supported in part by National Aeronautics and Space Agency through Grant NSG 330.

‡ Supported by the Air Force Office of Scientific Research.

[1] See, for example, W. Low, *Paramagnetic Resonance in Solids* (Academic Press Inc., New York, 1960).

[2] S. Foner and W. Low, Phys. Rev. **120**, 1585 (1960), have performed ESR experiments using pulsed magnetic fields of the order of 2 000 000 G. However, this method has the disadvantage of poor accuracy in the field measurements. Relaxation effects due to fast variation of the magnetic field as a function of time effect the intensity of the spectrum. It can be used, however, to some advantage in the measurement of relaxation times at high magnetic fields [L. Rimai (private communications)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

with $S=1$. When $E=0$ the initial separation of $|\pm1\rangle$ levels from the $|0\rangle$ level can be determined. When $E\neq0$ then the separation of one of the nearly degenerate levels from the $|0\rangle$ level can be measured.

Of greater interest are the rare-earth and uranium group elements. Here the various J levels are split into a number of Stark levels through the crystal field. In many cases, the total amount of the splitting of 5–10 Stark levels may be only of the order of a 100 cm^{-1}. There are levels separated by less than 20 cm^{-1}. The position of these levels can be estimated by diverse methods such as by measuring (1) the magnetic susceptibility or specific heat as a function of temperature, (2) the spin-lattice relaxation time as a function of the temperature and evaluation of the Orbach process,[3,4] (3) the fluorescence in the optical spectra, (4) the Faraday rotation, and (5) the intensity of the spin resonance of the excited state as a function of the temperature, provided this resonance can be observed in the temperature range where this level is still sufficiently populated. Most of these methods are not very accurate and permit the evaluation of the level only to an accuracy of about 1 cm^{-1}. Others do not work well in dilute salts which are used in spin resonance methods.

Since many of the g factors of rare-earth ions are large, one can bring levels, which in the absence of a magnetic field are far apart, close together and even cause them to cross.[5] Assuming a magnetic field of about 50 000 G and g factors of about 6 for the two levels, and operating K-band frequencies, one can determine energy levels separated by about 15 cm^{-1}.

The g factors in the rare-earth ions depend on the crystal-field parameters $A_n{}^m\langle r^n\rangle$. In general there are many more crystal-field parameters than the measured three or less g factors. Paramagnetic resonance measurements of these g factors on one single Kramers doublet does not permit the evaluation of these crystal-field parameters. These can be determined from the optical spectra which give the average value of the $A_n{}^m\langle r^n\rangle$ over all the optical levels. There are reasons to believe that the crystal-field parameters may depend to some extent on the various $J\Gamma$ levels because of the shielding and antishielding factors,[6,7] i.e., that $A_n{}^m\langle r^n\rangle_{\text{true}}(1-p_n{}^m(J,\Gamma))$, where $p_n{}^m$ is a shielding factor. It is, therefore, of interest to measure sufficient g factors and a separation of Stark levels of a given J level in order to determine the $A_n{}^m\langle r^n\rangle$ for this particular level and to compare it with the weighted average over all optical levels.

Separation of energy levels in the range of 1–20 cm^{-1} are also found in exchange coupled pairs of paramagnetic ions. For example the spectrum of exchange couplet Ir^{4+} in $K_2IrPtCl^6$ can be explained by the interaction spin Hamiltonian[8]

$$\mathcal{H}_{ij}=J_1\mathbf{S}^i\cdot\mathbf{S}^j+D(2S_z{}^iS_z{}^j-S_x{}^iS_x{}^j-S_y{}^iS_y{}^j)+E(S_x{}^iS_x{}^j-S_y{}^iS_y{}^j),$$
$$\text{with } S^i=S^j=\tfrac{1}{2}, \quad \text{and } J_1=8\ cm^{-1}. \quad (2)$$

The isotropic exchange parameter can be determined by one or more of the methods described above. Strong magnetic fields would permit a more accurate determination of these parameters.

Some of the g factors in paramagnetic resonance are smaller than one. An example of this is Sm^{3+}. A magnetic field of about 50 000 G and at K_μ band frequency would permit the determination of g factors as small as 0.2.

Finally, it would be very interesting to determine the spin-lattice relaxation in high magnetic fields, in particular, the magnetic-field dependence of the direct process, the Orbach process, and the Raman process both for Kramers and non-Kramers doublets.

We report here the techniques of using superconducting magnets for spin resonance experiments. We illustrate this by two types of experiments (a) the determination of the $\Gamma_8-\Gamma_7$ separation of Dy^{3+} in CaF_2 and (b) the determination of the small g factors of Sm^{3+} in CaF_2. Experiments as outlined in (c) and (d) are in progress.

EXPERIMENTAL DETAILS

We briefly describe here the experimental setup. A fuller description of the apparatus will be given elsewhere.

The experiments were made with a stabilized K_μ spectrometer operating between 15 000–18 000 Mc/sec and at liquid-helium temperature. The cavity and the stainless-steel waveguide section were enclosed in a 1-in. i.d. thin stainless-steel tubing. This tubing was inserted into the bore of the superconducting magnet. The modulation coils were wound around the stainless-steel tubing concentric with the cylindrical cavity. Fifty to 100 turns of thin pure copper wire were sufficient for a modulation depth of about 80 G. The heat dissipation at such modualtion depth was not excessive. About 8 liters of liquid helium were required to cool the superconducting magnet and waveguide section to fill the Dewar for a 3–4-h experiment.

The superconducting magnet was about 4 in. long

[3] R. Orbach, Proc. Roy. Soc. (London) **A264**, 458 (1961).
[4] C. B. P. Finn, R. Orbach, and W. P. Wolf, Proc. Phys. Soc. (London) **77**, 261 (1961).
[5] An interesting experiment is to measure the fluorescence or absorption of light when the energy levels cross. The polarization of the light or the angular distribution of the light would be expected to change. This experiment is very similar to the optical atomic level crossing experiments except that in the case of a crystal you have the advantage of the symmetry planes and axes. Such an experiment should permit the accurate determination of g factors and splittings of both of the ground state as well as optically excited states of transition group elements in crystals.
[6] R. E. Watson and A. J. Freeman, Phys. Rev. **133**, A1571 (1964).
[7] W. Low, Enrico Fermi Summer School, Varena, 1963 (to be published).
[8] J. H. E. Griffiths, J. Owen, J. G. Park, and M. F. Partridge, Proc. Roy. Soc. (London) **A250**, 84 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and i.d. slightly larger than 1 in.[9] The power supply was a standard stabilized supply. The magnetic field could be swept by putting a small dc error signal in the sensing elements of the power supply. The rate of sweep could be controlled by varying the magnitude of the error voltage. Sweeps of several minutes to several hours over the range of 46 000 G could be achieved without difficulty. A superconducting switch permitted the magnet to be operated in the persistent mode.

The measurement of the magnetic field presents a little difficulty. The inhomogeneity of this particular magnet over a sample volume of 0.02 cm^3 and a sample height of about 2 mm was of the order of 10–15 G at the center field of about 20 000 G. This inhomogeneity was estimated by putting small crystals with paramagnetic impurities of known linewidth and g factors. Calibrating the sweep, one could estimate the broadened linewidth in the superconducting magnet. The linewidth would probably be even larger for a proton resonance sample which has a larger volume and would not permit an accurate or easy determination of the magnetic field. (Since our spectrometer was inserted in the liquid-helium bath, a solid-state proton sample would have to be used. However, one could use also a liquid sample by using an inner Dewar so that the spectrometer and proton resonance could be operated at any temperature.) We have used a more primitive method. By placing various paramagnetic samples of known g factors in the magnetic field, we have made an experimental plot of the magnetic field H against the current I. It was found that the magnetic field was not quite linear with the current, particularly at high magnetic fields. There is also a noticeable hysteresis. At center band, near 20 000 G, and sweeping rapidly over the whole range back forth the hysteresis amounts to 100–150 G. In all measurements reported here were made with slow sweep rates and always in the same direction. The maximum error in the determination of the magnetic field is estimated to be about 0.2%.

Once a line was observed on the scope or by phase-sensitive detection the superconducting solenoid was switched to the persistent mode. In this case the line stayed at a constant magnetic field with only a very small drift.

Such a system, with a more homogenous magnet, could be used for magnet of frequency stabilization. It also permits the use of the continuous averaging technique[10] to greater advantage. In this technique the signal and the noise are fed into a multichannel analyzer. The signal is additive, i.e., proportional to the number n of sweeps across the line, whereas the noise is proportional to $n^{1/2}$, so that the signal to noise is proportional to $n^{1/2}$. This is also true for phase-sensitive detection. However, in practice it is found that low-frequency noise components are difficult to filter out, even with large time constants in the R.C. filter of lock-in detector. A system, such as a superconducting magnet, in which the magnetic field is constant and ripple is nearly absent, lends itself admirably to such a technique.

[9] Manufactured by Magnion, Inc., Cambridge, Massachusetts. The power supply is made by Kepko Model KS-36-30M.

[10] M. P. Klein and A. W. Barton, Jr., in *Proceedings of the First International Conference on Paramagnetic Resonance*, edited by W. Low (Academic Press Inc., New York, 1963), Vol. 2, p. 699.

THE SPECTRUM OF Dy^{3+} IN THE CUBIC FIELD OF CaF_2

The cubic field of CaF_2 splits the $J=15/2$ of the $^6H_{15/2}$ level into three Γ_8, one Γ_7 and one Γ_6. The ground state is a $\Gamma_8^{(1)}$ and the g factors within the Γ_8 quartet have recently been determined by Bierig and Weber[11] as 2.63 ± 0.05, 5.48 ± 0.15, and 14 ± 1. The next excited state Γ_7 has been measured by Low[12] as $g=7.47\pm0.03$ and by the above authors at 7.52 ± 0.05. These authors estimate from the temperature dependence of the intensity of the Γ_7 level to be at 8.5 ± 1.0 cm^{-1} above the Γ_8 level. All other levels are far removed. In addition Bierig and Weber observed many lines associated with tetragonal or trigonal symmetries and with g factors varying from 1.50 to 16.

This spectrum, therefore, has a wealth of lines which can be used for field calibration. We have confirmed many of the g factors determined by these authors (the high g factor is found to be 13.7 ± 0.3). In addition we observe lines at near 29 000 and 34 000 G when $H\|$ to the 100 axis. These two lines are identified to belong to the cubic transition before and after crossover from the

$$\Gamma_8:\ a_1|\pm15/2\rangle+a_2|\pm7/2\rangle+a_3|\mp1/2\rangle a_4|\mp9/2\rangle \text{ to the}$$

$$\Gamma_7:\ c_1|\pm13/2\rangle+c_2|\pm5/2\rangle+c_3|\mp3/2\rangle+d_4|\mp11/2\rangle \text{ levels.}$$

From these measurements and the known g factors of the Γ_8 and Γ_7 levels we evaluate the $\Gamma_8-\Gamma_7$ separation as $7.3_5\pm0.2$ cm^{-1}. The main errors arise from the additive errors of the g factors and the field calibration.

These data permit the evaluation of the crystal-field parameters. Writing in the notation of Lea *et al.*,[13]

$$\frac{A_4\langle r^4\rangle}{A_6\langle r^6\rangle}=\frac{x}{1-|x|}\frac{F(6)}{F(4)},$$

where $A_m\langle r^n\rangle$ are the crystal-field parameter, $F(4)$ and $F(6)$ multiplication factors, and we find $x=0.59\pm0.01$. From this we find

$$A_4\langle r^4\rangle=-242\pm20 \text{ cm}^{-1},$$

$$A_6\langle r^6\rangle=41\pm5 \text{ cm}^{-1}.$$

[11] R. W. Bierig and M. J. Weber, Phys. Rev. 132, 164 (1963).

[12] W. Low, Proc. Phys. Soc. (London) 76, 307 (1960).

[13] K. R. Lea, M. J. Leask, and W. P. Wolf, Phys. Chem. Solids 23, 1381 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In these calculations we have not taken into account the admixtures from the $J=13/2$ level at about 3500 cm^{-1}.

THE SPECTRUM OF Sm^{3+} IN THE TEGRAGONAL FIELD OF CaF_2

The Landé g factor of the ${}^6H_{5/2}$ level is small, only 2/7, and one would expect that the spectroscopic splitting factor of the Sm^{3+} in crystals would be small as well.

In the cubic field the $J=5/2$ levels is split into a Γ_8 and Γ_7. The Γ_8 level in the absence of admixture of higher excited states gives rise to transitions at $g=22/21$, 2/3, and 2/7 with relative intensity of 25/36, 2/3, and 9/8, respectively, when $H\|$ (100). With the magnetic field parallel to the (111) direction the calculated g factors are 16/21, 2/3, and 2/12 with relative intensities of 16/9, 4/9, 25/18. However, the $J=7/2$ is only at about 1000 cm^{-1} followed by $J=9/2$ near 2000 cm^{-1} and one expects that the various levels would be admixed considerably, reducing in general the g factors from the values given above. In the presence of an axial field the Γ_8 level splits into two doublets. Rabbiner[14] has observed the fluorescence spectrum of Sm^{3+} supposedly in an axial field and has analyzed the spectrum. The interpretation, however, should be treated with caution because of the complexity of this spectrum.

We have looked for the resonance spectrum of two CaF_2 with nominal concentration of 0.05 and 0.1%. One set of crystals was nearly colourless supposedly contained mainly trivalent Sm^{3+}, and the other set of deep green color and containing a mixture of Sm^{3+} and Sm^{2+}. The spectrum of the green crystal was very rich in lines whereas the other showed only a few lines. The lines common to both samples could be shown to have tetragonal point symmetry with

$$g_{\|}=0.907\pm0.01,$$
$$g_{\perp}=0.544\pm0.15.$$

[14] N. Rabbiner, Phys. Rev. **130**, 507 (1963).

[*Note added in proof.* M. J. Weber and R. W. Bierig have observed a similar spectrum with

$$g_{\|}=0.93\pm0.04$$
$$g_{\perp}=0.57\pm0.05$$

(private communication).]

In addition there were many other lines (apart of hyperfine structure from the Sm^{147} and Sm^{149}, both with spins $I=\frac{7}{2}$). These lines have not been yet interpreted, some of which no doubt are of cubic and trigonal symmetry. The presence of the rich hyperfine structure complicated the identification of the spectra. For example along the (111) direction, lines were found with $g=0.893$, 0.87, 0.70, 0.67 (tetragonal), 0.62 and 0.415. (It is likely that line at 0.415 belongs to the cubic set. Possibly the line at $g=0.89$ or 0.87 may also belong to the cubic set. However, this needs further confirmation.) It is of interest to note that there are so many more lines in the mixed $Sm^{2+}-Sm^{3+}$ crystal. Possibly some of the Sm^{2+} which presumably is in the cubic site oxidizes later to Sm^{3+}.

The above g factor can be interpreted to arise from a wave function of the $a|\pm5/2\rangle+b|\mp3/2\rangle$ with a ~0.86. This is indicative of considerable tetragonal admixture and is consistent with similar results found for[15,16] Nd^{3+} and for Ce^{3+}.[17]

ACKNOWLEDGMENTS

The author is grateful to Bruce Montgomery for permission to use the superconducting magnet and for discussions regarding the operation of this magnet and to James Casteris for technical assistance with the measurements.

[15] B. Bleaney, P. M. Llewellyn, and D. A. Jones, Proc. Phys. Soc. (London) **B69**, 858 (1956).
[16] A. Vincow and W. Low, Phys. Rev. **133**, 1390 (1961).
[17] J. M. Baker, W. Hayes, and D. A. Jones, Proc. Phys. Soc. (London) **73**, 942 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Paramagnetic Resonance of Yb^{3+} in $CaWO_4$†

U. RANON
Israel Atomic Energy Commission, Soreq Research Establishment, Yavne, Israel
AND
V. VOLTERRA
Department of Physics, The Hebrew University, Jerusalem, Israel
(Received 31 January 1964)

Paramagnetic resonance of Yb^{3+} has been observed in single crystals of $CaWO_4$ at 20°K and 3 cm. A tetragonal as well as two rhombic spectra were found. The tetragonal spectrum is fitted to an axial spin Hamiltonian with $g_{\parallel}=1.058\pm0.002$, $g_{\perp}=3.920\pm0.005$, $|^{173}A|=(77\pm1)10^{-4}$ cm^{-1}, $|^{173}B|=(285\pm1)10^{-4}$ cm^{-1}, $|^{173}P|<2.10^{-3}$ cm^{-1}, $|^{171}A|=(279\pm3)10^{-4}$ cm^{-1}, $|^{171}B|=(1034\pm3)10^{-4}$ cm^{-1}. The rhombic spectra have the y and z axes in the {110} planes. One has the z axis at $6°\pm1°$ to the crystallographic c axis and $g_x=3.012\pm0.005$, $g_y=4.788\pm0.005$, $g_z=0.975\pm0.003$. The other has the z axis at $7°\pm1°$ to the c axis and $g_x=4.155\pm0.005$, $g_y=3.704\pm0.005$, $g_z=1.155\pm0.003$. The spectra arise from Yb^{3+} ions at Ca^{2+} sites and the rhombic symmetry is caused by charge compensation at next-nearest-neighbor Ca^{2+} sites. The possibility of W^{6+} substitution is discussed.

I. INTRODUCTION

CRYSTALS of the $CaWO_4$ structure (sheelite), doped with rare-earth ions have been used extensively in the past few years as solid-state lasers. This fact stimulated the optical as well as paramagnetic resonance studies of these crystals.[1,2]

Paramagnetic resonance in $CaWO_4$ has been reported on the trivalent ions gadolinium,[3] terbium,[4] and neodymium.[5,6] In all these cases the symmetry of the crystalline field at the sites of the paramagnetic ions was found to be tetragonal. Recently, orthorhombic spectra of Fe^{3+} and Nd^{3+} have been reported.[6] These spectra were attributed to substitutional sites of W^{6+} ions in the WO_4^{2-} complex.

The purpose of this paper is to report on the investigation of Yb^{3+} in $CaWO_4$. A tetragonal as well as two rhombic spectra have been found and the results cast some doubt on the attribution of rhombic spectra to the W^{6+} sites.

II. THE STRUCTURE OF $CaWO_4$

$CaWO_4$ (sheelite) is a tetragonal body-centered crystal which belongs to the C_{4h} point group. The point symmetry at the sites of the Ca^{2+} ions is S_4. The structure is shown in Fig. 1. The distorted oxygen tetrahedra which surround the W^{6+} ions, are represented schematically.

III. THEORY

The ion Yb^{3+} has $4f^{13}$ electrons and the ground term of the free ion is $^2F_{7/2}$. In the rare earths, the interaction with the crystal field is usually much weaker than the spin-orbit coupling and therefore J is a good quantum number. In crystal fields of lower than cubic symmetry the J manifold of the ground state of the free Yb^{3+} ion, splits into four Kramers doublets. Paramagnetic resonance is observable at liquid-hydrogen temperature where only the lowest Kramers doublet is populated.

IV. RESULTS

Two crystals of $CaWO_4$, doped with Yb^{3+}, were investigated by paramagnetic resonance at 3 cm and 20°K. They were obtained from Isomet Corporation and Semi-Elements Inc. The concentration of Yb^{3+} in the former was nominally 0.01% and higher in the latter. An intense tetragonal spectrum was observed in the two crystals and the hyperfine structure of the

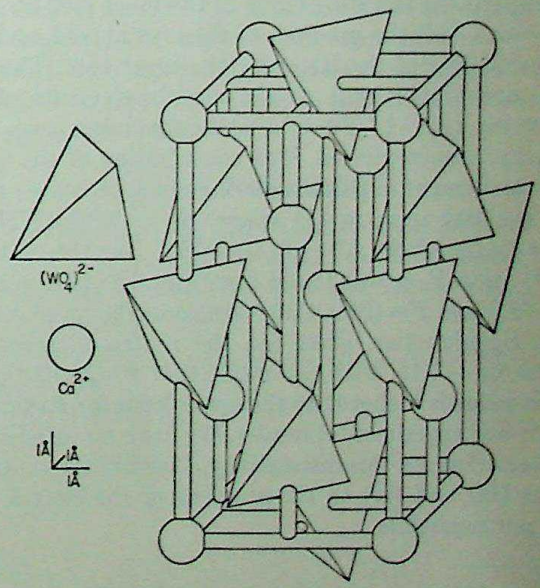


FIG. 1. Unit cell of $CaWO_4$. For clarity the oxygen tetrahedra are presented schematically.

† Supported in part by the U. S. Air Force through the European Office, under Grant No. AF-EOAR 63-64.

[1] L. F. Johnson, J. Appl. Phys. 34, 897 (1963).
[2] L. F. Johnson and R. A. Thomas, Phys. Rev. 131, 2038 (1963).
[3] C. F. Hempstead and K. D. Bowers, Phys. Rev. 118, 131 (1960).
[4] P. A. Forrester and C. F. Hempstead, Phys. Rev. 126, 923 (1962).
[5] U. Ranon, Phys. Letters 8, 154 (1964).
[6] R. W. Kedzie and K. Kestigian, Appl. Phys. Letters 3, 86 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

isotopes Yb 171($I=\frac{1}{2}$) and Yb 173($I=\frac{5}{2}$) was clearly resolved. Two spectra of rhombic symmetry which were much weaker, were also shown by each of the two crystals.

The Tetragonal Spectrum

The tetragonal spectrum fits the spin Hamiltonian,

$$\mathcal{H}=g_{11}\beta H_z S_z+g_1\beta(H_x S_x+H_y S_y)+AS_zI_z+B(S_xI_x+S_yI_y)+P(I_z^2-1/3I(I+1)), \quad (1)$$

where the z axis is parallel to the crystallographic c axis and the parameters are

$$g_{11}=1.058\pm0.002,$$
$$|^{173}A|=(77\pm1)10^{-4}\ \text{cm}^{-1},$$
$$|^{173}B|=(285\pm1)10^{-4}\ \text{cm}^{-1},$$
$$|^{173}P|<2.10^{-3}\ \text{cm}^{-1},$$
$$g_1=3.920\pm0.005,$$
$$|^{171}A|=(279\pm3)10^{-4}\ \text{cm}^{-1},$$
$$|^{171}B|=(1034\pm3)10^{-4}\ \text{cm}^{-1}.$$

These results give $|g_{11}A/g_1B|=1$ within the experimental error, and therefore the admixture of $J=\frac{5}{2}$ into the ground state is negligible.[7] Accordingly, the Kramers doublet can be described by an eigenfunction within the $J=\frac{7}{2}$ manifold. In a tetragonal field this eigenfunction may be of the form $a|\pm\frac{7}{2}\rangle+b|\mp\frac{1}{2}\rangle$ or $c|\pm\frac{5}{2}\rangle+d|\mp\frac{3}{2}\rangle$. Only the latter expression is compatible with the measured g factors, giving $g_{11}=1.052$ and $g_1=3.957$ when $c=0.700$ and $d=0.714$.

The gradient of the electric field at the nucleus is determined by four contributions: (i) that of the $4f$ electrons, (ii) the shielding effect of the inner core on these electrons, (iii) the gradient of the crystal field, and (iv) the shielding of the latter by the inner core. The first two are proportional to the matrix elements of the operator $(J_z^2-\frac{1}{3}J(J+1))$, while the other two have usually been neglected. The contribution of the quadrupole moment in paramagnetic resonance is then given by the last term in (1) where[7] $P=-9e^2Q/4I(2I+1)\times\langle r^{-3}\rangle\langle J\|\alpha\|J\rangle\langle\Psi|J_z^2-\frac{1}{3}J(J+1)|\Psi\rangle$. In the case of Yb^{3+} in CaF_2 it was found that[8] $\langle\Psi|J_z^2-\frac{1}{3}J(J+1)|\Psi\rangle=-3.9$ and $P=(86\pm4)10^{-4}\ \text{cm}^{-1}$. In the present case, the eigenfunction gives for the matrix element the value of -1.16, with the result that $P=25.10^{-4}\ \text{cm}^{-1}$. This value is higher than the upper limit for P obtained from our experimental results. We have no satisfactory explanation for this discrepancy. Possibly the contributions (iii) and (iv) to the gradient of the electric field are not negligible.

[7] R. J. Elliott and K. W. H. Stevens, Proc. Roy. Soc. (London) A218, 553 (1953).
[8] W. Low and U. Rosenberger, Compt. Rend. 254, 1771 (1962).

The Rhombic Spectra

The rhombic spectra can be fitted to the spin Hamiltonian

$$\mathcal{H}=\mathbf{H}.g.\mathbf{S}+\mathbf{S}.A.\mathbf{I},$$

where g and A are tensors with principal values g_x, g_y, g_z, and A_x, A_y, A_z, respectively.

Each of the rhombic spectra shows the existence of four inequivalent sites. For both spectra, one of the orthorhombic axes is perpendicular to the c direction, and we call it the y axis. Two sites of each spectrum have y along the [110] direction whereas the other two have it along the $[1\bar{1}0]$ direction. The other two axes lie then in the (110) and $(1\bar{1}0)$ planes, respectively. The sites of the more intense of the two rhombic spectra, which we shall call spectrum one, have the z axis at an angle of $6°\pm1°$ to c, while the sites which belong to the weaker spectrum (number two), have the z axis at an angle of $7°\pm1°$ to c.

Two orientations of z with respect to c are possible in each of the (110) and $(1\bar{1}0)$ planes, giving rise to four inequivalent sites which are interchangeable in pairs by a rotation of 90° about the c axis.

Spectrum number one has the parameters $g_x=3.012\pm0.005$, $g_y=4.788\pm0.005$, and $g_z=0.975\pm0.003$.

Spectrum number two has the parameters $g_x=4.155\pm0.005$, $g_y=3.704\pm0.005$ and $g_z=1.155\pm0.003$. The intensity of the even isotope lines in spectrum number one, was approximately 5% of that in the tetragonal spectrum, and the intensity in spectrum number two was about 2% of that of the tetragonal lines. On account of the four inequivalent sites for each rhombic spectrum it is found that about 16% of the Yb^{3+} ions occupy rhombic sites which correspond to spectrum one, and about 6% occupy sites corresponding to spectrum two.

The relative abundance of the two odd isotopes, 171 and 173, of ytterbium is 14% and 17%, respectively and therefore the hyperfine lines of the rhombic spectra were masked by the hyperfine lines of the tetragonal spectrum in such a way that only a few of the former were observable. Therefore the parameters of A could not be evaluated.

From the results of the tetragonal spectrum we assume that in the rhombic sites also, the admixture of $J=\frac{5}{2}$ is negligible. Since now the symmetry is rhombic, the eigenfunctions are of the form $a|\pm\frac{7}{2}\rangle+b|\mp\frac{1}{2}\rangle+c|\mp\frac{5}{2}\rangle+d|\pm\frac{3}{2}\rangle$. The three g values g_x, g_y, g_z and the normalization of the function give four quadratic equations in the four coefficients (a,b,c,d). In this case therefore the existence of a solution of the four equations does not imply that the admixture of $J=\frac{5}{2}$ is negligible, as opposed to the tetragonal case where there are three equations in two unknowns.

From the g values it can be seen that the deviation of the rhombic field from the tetragonal is not large. Therefore, we have looked for a function which is close to

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that in the tetragonal case, obtaining for spectrum 1, $0.160|\pm\frac{7}{2}\rangle-0.262|\mp\frac{1}{2}\rangle+0.678|\mp\frac{5}{2}\rangle+0.680|\pm\frac{3}{2}\rangle$ and for spectrum number 2,

$$-0.093|\pm\tfrac{7}{2}\rangle+0.099|\mp\tfrac{1}{2}\rangle+0.708|\mp\tfrac{5}{2}\rangle+0.693|\pm\tfrac{3}{2}\rangle.$$

V. DISCUSSION

The tetragonal spectrum is usually interpreted as being due to ions at substitutional sites of Ca^{2+}. Kedzie and Kestigian[6] explained their rhombic spectra of Fe^{3+} and Nd^{3+} as being due to ions which occupy tungsten sites in the WO_4^{2-} complex. The distortion from tetragonal symmetry is explained by them in these cases by assuming a deformation of one of the oxygen bonds along the bond direction in the substituted WO_4 complex, giving rise to four inequivalent sites.

We disagree with this interpretation of the rhombic spectra for the following reasons.

(1) It seems unlikely that trivalent ions, such as the rare-earth ions should substitute for the hexavalent tungsten, which is part of a covalent complex. Furthermore, the ionic radii of the rare-earth ions are much larger than the radius of W^{6+}.

(2) The W-O bonds in the WO_4^{2-} complex are not situated in any crystallographic plane.[9] A deformation of such a bond, therefore, cannot cause a rhombic distortion with axes lying in the (110) plane, whereas our investigations, as also those of Kedzie and Kestigian, show that such rhombic distortion does actually occur.

(3) There seems to be no reason why the deformation of the complex—which is situated in a tetragonal environment—will be such as to prefer one single direction (namely, a bond direction) which does not coincide with the tetragonal axis.

(4) In our case, *two* rhombic spectra are observed. Even if we assume the W^{6+} substitution, only one of these spectra could be explained. Furthermore, the crystal field at the tungsten sites will be completely different from the tetragonal field, while our results show only a small deformation of the tetragonal field at both rhombic sites. This is evident from the fact that the g values of the rhombic spectra do not differ considerably from those of the tetragonal spectra and that the trace of the g tensor is almost the same in the three cases.

Therefore we assign the rhombic spectra of Yb^{3+} to ions at Ca^{2+} sites where the deformation of the tetragonal field is caused by charge compensation. It is well known that in crystals like CaF_2, where also trivalent ions substitute Ca^{2+}, different charge compensations are possible and these give rise to different local symmetries of the crystal field.

In $CaWO_4$ two forms of charge compensation suggest themselves. These are: (i) Ca^{2+} vacancies, and (ii) substitutional Ca^{2+} sites occupied by monovalent cations when the crystals are grown with such ions.

The nearest-neighbor Ca^{2+} sites lie in the (100) and (010) planes while the orthorhombic axes lie in the (110) planes. The latter is consistent with charge compensation at next-nearest-neighbor Ca^{2+} sites. Cases where charge compensation is affected at more distant sites rather than at nearest-neighbor sites, were observed in CaF_2[10] as well as in SrF_2 and BaF_2.[11]

Hempstead and Bowers[3] and Kedzie and Kestigian[6] found that Mn^{2+} (which does not require charge compensation) is situated exclusively in a tetragonal field. The latter authors found that the isoelectronic Fe^{3+} (which does require charge compensation) shows only an orthorhombic spectrum.

The fact that the spectra of Mn^{2+} and Fe^{3+} are completely different is not, in our opinion, a conclusive indication that Fe^{3+} does not occupy Ca^{2+} sites, as suggested by Kedzie and Kestigian. As these authors pointed out, the spectrum of Fe^{3+} can only be explained by assuming a strong axial field where the Stark splitting is larger than the Zeeman splitting. The following consideration shows that this may occur at the Ca^{2+} sites; it is well known that the Stark splitting in Fe^{3+} is at least one order of magnitude larger than in Mn^{2+}. The splitting[3] between the closest doublets of Mn^{2+} at the Ca^{2+} sites is 0.0275 cm^{-1}. Thus the splitting in Fe^{3+} should be of the order of the Zeeman splitting, or larger.

In view of the present results and the uncertainty in the assignment of the rhombic sites, it seems worthwhile that further investigations be undertaken of other rare-earth ions in $CaWO_4$ as well as in other crystals having the sheelite structure.

ACKNOWLEDGMENTS

The authors are grateful to Zvi Kalman for the x-ray determination of the crystal axes. Stimulating discussions with Professor W. Low and Dr. S. Yatsiv are gratefully acknowledged.

[9] R. W. G. Wyckoff, *Crystal Structures* (Interscience Publishers Inc., New York 1957), Vol. 2, Chap. 8.

[10] E. Mahlab, W. Low, V. Volterra, and A. Yariv, Phys. Rev. 131, 920 (1963).

[11] U. Ranon and A. Yaniv, Phys. Letters 9, 17 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Frequency Spectra of Body-Centered Cubic Lattices*

B. C. CLARK
General Motors Research Laboratories, Warren, Michigan

AND

D. C. GAZIS
IBM Research Center, Yorktown Heights, New York

AND

R. F. WALLIS
U. S. Naval Research Laboratory, Washington, D. C.

(Received 27 January 1964)

A Born-von Karman model has been used to determine the frequency spectrum for a body-centered cubic lattice. The model used contains noncentral angular stiffness forces as well as central forces between nearest and next-nearest neighbors. Frequency distributions have been calculated for vanadium and compared with experimental frequency distributions available from slow neutron experiments. Qualitative agreement was found between the shapes of the experimental and theoretical distributions. The calculated maximum frequency for vanadium was within 2% of that derived from the experimental frequency distribution. Dispersion curves for iron were also calculated, and were found to be in good agreement with experiment.

INTRODUCTION

THE frequency distribution of a body-centered cubic lattice of identical particles has been investigated using the Born-von Karman model of a crystal lattice. The interatomic forces used include noncentral forces of the type introduced by Gazis, Herman, and Wallis[1] in treating surface waves in a simple cubic crystal. These noncentral forces arise from the resistance to deformation of certain angles formed by three lattice points. Also included are central forces between nearest and next-nearest neighbors. Noncentral forces have been employed in recent calculations by Hendricks, Riser, and Clark,[2] which appeared after our work was essentially complete; however, the noncentral forces considered by them are different from those used here.

One of the primary aims of the investigation was to examine the effect of the inclusion of noncentral forces on the frequency distribution. Comparison between the central-force model and the model employed here was made for iron and for vanadium which is one of the more anisotropic body-centered cubic crystals. The central-force calculations for vanadium were available from the work of Clark.[3] The calculations with vanadium were also compared with the frequency distributions obtained by the slow neutron experiments of Stewart and Brockhouse,[4] and Eisenhauer *et al.*[5] Dispersion curves along specified directions were calculated for iron and compared with the experimentally determined dispersion curves of Low.[6]

LATTICE EQUATIONS OF MOTION

Although the equations of motion for central forces appear elsewhere, for example, in DeLaunay,[7] they will be repeated here in order to provide a continuity of notation. Taking a displacement vector with components u, v, and w in the x, y, and z directions, respectively, the force in the x direction on atom (j, m, n) due to the central-force interactions of nearest neighbors is given by

$$F_{x1}=\alpha[U-8u_{j,m,n}+V+W], \tag{1}$$

where

$$\begin{aligned}U=&(u_{j+1,m+1,n+1}+u_{j+1,m+1,n-1}+u_{j+1,m-1,n+1}\\&+u_{j+1,m-1,n-1}+u_{j-1,m+1,n+1}+u_{j-1,m+1,n-1}\\&+u_{j-1,m-1,n+1}+u_{j-1,n-1,m-1}),\\V=&(v_{j+1,m+1,n+1}+v_{j+1,m+1,n-1}-v_{j+1,m-1,n+1}\\&-v_{j+1,m-1,n-1}-v_{j-1,m+1,n+1}-v_{j-1,m+1,n-1}\\&+v_{j-1,m-1,n+1}+v_{j-1,m-1,n-1}),\\W=&(w_{j+1,m+1,n+1}-w_{j+1,m+1,n-1}+w_{j+1,m-1,n+1}\\&-w_{j+1,m-1,n-1}-w_{j-1,m+1,n+1}+w_{j-1,m+1,n-1}\\&-w_{j-1,m-1,n+1}+w_{j-1,m-1,n-1}),\end{aligned} \tag{2}$$

where α is the nearest-neighbor force constant. Similarly, the central-force interaction between next-nearest neighbors gives rise to a force on the atom (j,m,n) in the x direction given by

$$F_{x2}=\beta[u_{j+2,m,n}+u_{j-2,m,n}-2u_{j,m,n}], \tag{3}$$

where β is the next-nearest-neighbor force constant.

---

* A part of this work was submitted by one of the authors (B. C. C.) in partial fulfillment for the degree Master of Science at Kansas State University.

[1] D. C. Gazis, R. Herman, and R. F. Wallis, Phys. Rev. **119**, 533 (1960).

[2] J. B. Hendricks, H. B. Riser, and C. B. Clark, Phys. Rev. **130**, 1377 (1963).

[3] C. B. Clark, J. Grad. Res. Center **29**, 10 (1961).

[4] A. T. Stewart and B. N. Brockhouse, Rev. Mod. Phys. **30**, 250 (1958).

[5] C. M. Eisenhauer, I. Pelah, D. J. Hughes, and H. Palevsky, Phys. Rev. **109**, 1046 (1958).

[6] G. G. E. Low, Proc. Phys. Soc. (London) **79**, 479 (1962).

[7] J. DeLaunay, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1956), Vol. 2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

For the noncentral forces examined here it is assumed that a change in angle is related to a change in potential energy by a Hooke's law force constant. If V is the potential energy and $\delta\phi$ a change in angle, one obtains

$$V=\tfrac{1}{2}\gamma(\delta\phi)^2, \tag{4}$$

where γ is the force constant. In this treatment two general types of angles were considered; 36 angles of the type ϕ_1 in Fig. 1 associated with a force constant γ_1, and 72 angles of the type ϕ_2 in Fig. 1 associated with a force constant γ_2. The expansion of the potential energy is quite tedious but straightforward, and the total contribution from the angles considered results in a force on the atom of interest, in the x direction, given by

$$\begin{aligned}F_{x3}=&\gamma_1[-20u_{j,m,n}+2U-V-W+4(u_{j+2,m,n}+u_{j-2,m,n})\\&-(u_{j,m+2,n}+u_{j,m-2,n}+u_{j,m,n+2}+u_{j,m,n-2})]\\&-\gamma_2[-30u_{j,m,n}+3U+\tfrac{3}{2}(V+W)\\&+\tfrac{3}{2}(u_{j,m+2,n}+u_{j,m-2,n}+u_{j,m,n+2}+u_{j,m,n-2})],\end{aligned} \tag{5}$$

where U, V, and W are given by Eq. (2). These noncentral forces are invariant under rigid-body rotations.

The equation for the force in the x direction on the atom (j,m,n) due to the central and noncentral forces considered here may now be written as

$$\begin{aligned}F_x=&m\ddot{u}_{j,m,n}\\=&-u_{j,m,n}[8\alpha+2\beta+20\gamma_1+30\gamma_2]\\&+(\beta+4\gamma_1)(u_{j+2,m,n}+u_{j-2,m,n})+(-\gamma_1+\tfrac{3}{2}\gamma_2)\\&\times(u_{j,m+2,n}+u_{j,m-2,n}+u_{j,m,n+2}+u_{j,m,n-2})\\&+U(\alpha+2\gamma_1+3\gamma_2)+(V+W)(\alpha-\gamma_1+\tfrac{3}{2}\gamma_2),\end{aligned} \tag{6}$$

where U, V, and W are given as in Eq. (5). The forces in the y and z directions on the central atom are found by a circular permutation of u, v, and w, as well as the increments of their indices j, m, n.

SOLUTION OF THE EQUATIONS OF MOTION

The solutions of the equations of motion are assumed to be plane waves, where the usual cyclic boundary conditions of Born[8] are applied. The solutions for the motion of, say the nth, atom are written as

$$\begin{aligned}u_n&=X\exp 2\pi i(\nu t+\mathbf{k}\cdot\mathbf{r}_n),\\v_n&=Y\exp 2\pi i(\nu t+\mathbf{k}\cdot\mathbf{r}_n),\\w_n&=Z\exp 2\pi i(\nu t+\mathbf{k}\cdot\mathbf{r}_n).\end{aligned}$$

Here X, Y, and Z are the components of the plane-wave amplitude, and $\mathbf{k}$ is a propagation vector in the reciprocal lattice whose magnitude is the reciprocal of the wavelength. The values of $\mathbf{k}$ are restricted to the first Brillouin zone, as discussed by Brillouin.[9] Substitution of these solutions into the equations of motion gives three homogeneous linear equations in X, Y, and Z.

[8] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon Press, Oxford, 1954).

[9] L. Brillouin, *Wave Propagation in Periodic Structures* (Dover Publications, New York, 1953).

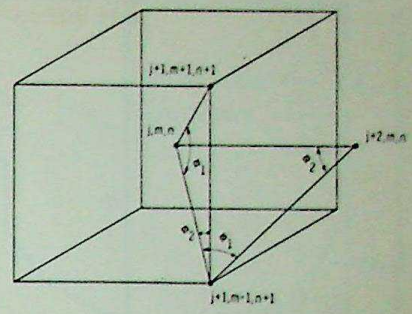


FIG. 1. Body-centered cubic lattice showing primitive (ϕ_1) and secondary (ϕ_2) angles.

A nontrivial solution requires that the determinant of the coefficients of X, Y, and Z vanish. This condition yields the secular equation

$$\begin{vmatrix}A_1-m\omega^2 & B & C\\ B & A_2-m\omega^2 & D\\ C & D & A_3-m\omega^2\end{vmatrix}=0, \tag{7}$$

where

$$\begin{aligned}A_i=&(8\alpha+16\gamma_1+24\gamma_2)(1-\cos\Phi_1\cos\Phi_2\cos\Phi_3)\\&+4\beta\sin^2\Phi_i-2\gamma_1(5\cos 2\Phi_i-\cos 2\Phi_1-\cos 2\Phi_2\\&-\cos 2\Phi_3-2)+3\gamma_2(2+\cos 2\Phi_i-\cos 2\Phi_1\\&-\cos 2\Phi_2-\cos 2\Phi_3),\end{aligned} \tag{8}$$

$$B=8(\alpha-\gamma_1+\tfrac{3}{2}\gamma_2)\cos\Phi_3\sin\Phi_1\sin\Phi_2,$$

$$C=8(\alpha-\gamma_1+\tfrac{3}{2}\gamma_2)\cos\Phi_2\sin\Phi_1\sin\Phi_3,$$

$$D=8(\alpha-\gamma_1+\tfrac{3}{2}\gamma_2)\cos\Phi_1\sin\Phi_2\sin\Phi_3.$$

In Eq. (8), $\Phi_i=\pi ak_i$ $(i=1,2,3)$ where a is the distance between next-nearest neighbors for the body-centered cubic lattice.

Symmetry considerations allow a restriction of the values of Φ_i to points within 1/48 of the first Brillouin zone. The limits on the Φ_i are then given by $0\leq\Phi_1\leq\frac{1}{2}\pi$, $0\leq\Phi_2\leq\frac{1}{2}\pi$, and $0\leq\Phi_3\leq\pi$. This reduces the values of Φ_i to those within the region in Φ space bounded by the planes $\Phi_1=\Phi_2$, $\Phi_2=0$, $\Phi_1=\Phi_3$ and $\Phi_1+\Phi_3=\pi$.

For a cubic equation of the type resulting from the expansion of Eq. (7) Blackman[10] has shown that maxima in $m\omega^2$ can occur only when Φ_1, Φ_2, Φ_3 have some combination of the values 0, $\frac{1}{2}\pi$, π. The point $\Phi_1=\Phi_2=\Phi_3=0$ is excluded as $m\omega^2=0$ at this point. Substitution of those values of Φ_1, Φ_2, Φ_3 which are within the first 1/48 Brillouin zone showed that in most cases considered the maximum occurs at $(\frac{1}{2}\pi,0,\frac{1}{2}\pi)$. This maximum frequency may be found directly from Eq. (7) and is

$$m\omega_{\max}^2=4(4\alpha+\beta+5\gamma_1)+42\gamma_2. \tag{9}$$

It was found in actual calculations that the position of the maximum could be shifted to $(0,0,\pi)$ when certain values of the elastic constants, such as those for sodium, given by Quimby and Siegell,[11] were assumed. In this case the maximum frequency is given by

$$m\omega_{\max}^2=16(\alpha+2\gamma_1+3\gamma_2). \tag{9a}$$

[10] M. Blackman, Proc. Roy. Soc. (London) **A148**, 365 (1935).

[11] S. L. Quimby and S. Siegell, Phys. Rev. **54**, 293 (1938).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

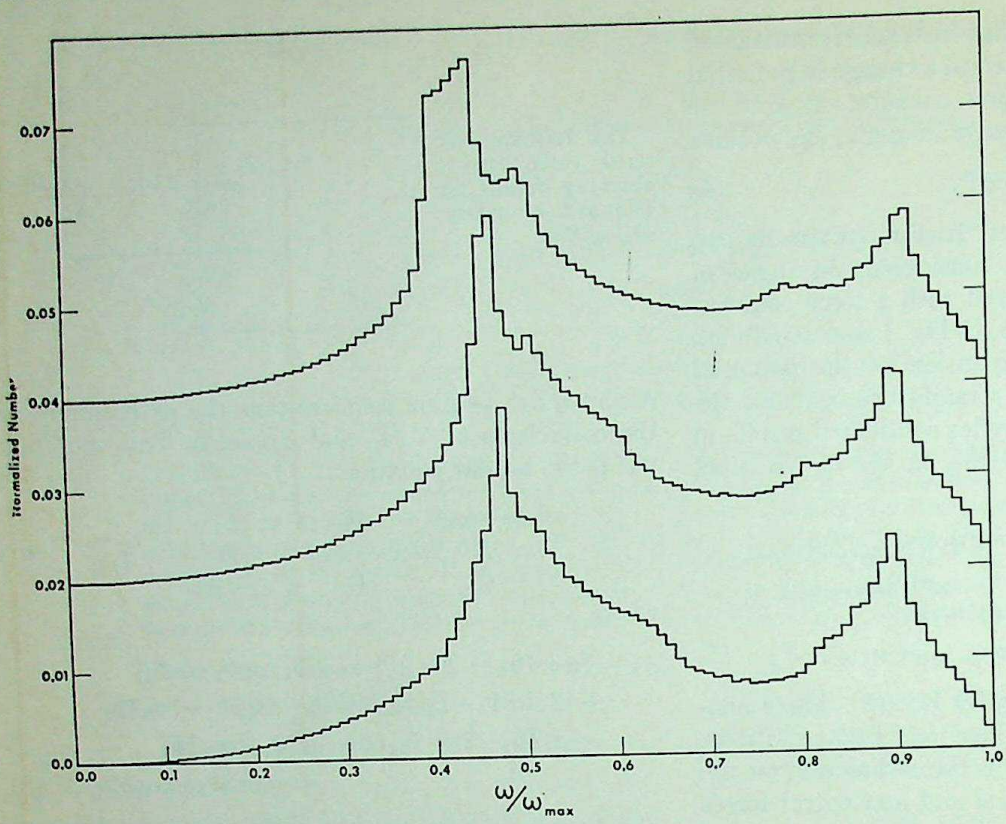


FIG. 2. Three calculated frequency histograms for vanadium with $\gamma_2/\gamma_1=0.0$ (upper curve), $\gamma_2/\gamma_1=0.5$ (middle curve), and $\gamma_2/\gamma_1=1.0$ (bottom curve). *Note:* For display purposes, while the scale of the ordinate is the same for all three curves, the origin for the middle curve is at 0.02 and is at 0.04 for the upper curve.

RELATIONSHIPS BETWEEN ELASTIC CONSTANTS AND LATTICE FORCE CONSTANTS

In the limit of long wavelengths, when the wavelength is much greater than the lattice spacing, the plane waves are effectively propagating as in a continuous medium. This allows the identification of the lattice force constants with the elastic constants of the medium found from measurements of the velocity of sound. These relationships are obtained by expanding the displacement components such as $u_{j+p,m+q,n+r}$ in a Taylor series about $u_{j,m,n}$, substituting into the equation of motion and matching the coefficients of the second order derivatives with those of the continuum equations of motion. The results are:

$$\begin{aligned} ac_{11}&=2(\alpha+\beta+6\gamma_1+3\gamma_2),\\ ac_{12}&=2(\alpha-3\gamma_1+\tfrac{3}{2}\gamma_2),\\ ac_{44}&=2(\alpha+\gamma_1+\tfrac{9}{2}\gamma_2). \end{aligned} \tag{10}$$

If $\gamma_2=0$, the above three equations determine the lattice force constants uniquely; otherwise, the force constants are determined assuming, for example, a given ratio of γ_2/γ_1. The value γ_2/γ_1 is used as a parameter in the sequel. The elastic constants for vanadium are those of Alers.[12] It may be observed that the maximum frequency when it occurs at $(\frac{1}{2}\pi,0,\frac{1}{2}\pi)$, as is usual, is a function of the elastic constants only and is

$$m\omega_{\max}{}^2=2a(c_{12}+2c_{44}+c_{11}). \tag{11}$$

However, when the maximum frequency occurs elsewhere it is also a function of γ_2/γ_1.

[12] G. A. Alers, Phys. Rev. 119, 1532 (1960).

CALCULATION OF THE FREQUENCY DISTRIBUTION

In order to obtain the true distribution of frequencies in a macroscopic crystal Eq. (7) would have to be solved for a very large number of points within the first Brillouin zone, equal to the number of degrees of freedom of the particles in the crystal. A systematic method was used for choosing a relatively small number of solution points within the first zone. It is assumed that solution points are uniformly distributed such that an allowed point exists within a volume element in Φ space, and further that the frequencies at all points within the element are equal to the value of the frequency at the center of the element. The values of Φ_i were chosen in a regular manner by stepping along the axes in intervals of 0.01π. Thus the volume elements were cubes of volume $10^{-6}\,\pi^3$.

Some of the solution points were on the bounding planes and were not, of course, centers of a cube lying entirely within the first Brillouin zone. There was some question of the desirability of including these points without properly weighing their contribution to the frequency distribution in the first zone. A numerical investigation of the most extreme case, that of the frequency distribution obtained with all points on the bounding planes included with the same weight as the interior points, showed that the resulting frequency distribution was essentially the same as the distribution with no boundary points included. Due to this result only points in the interior of the first 1/48th Brillouin zone were finally chosen. These considerations resulted in the inclusion of approximately 39 000 solution points in the 1/48th zone. Histograms representing the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

frequency distribution were prepared from the calculated frequencies by finding the number of frequencies contained in a given frequency interval. The frequency interval chosen was 1.0% of the maximum frequency.

The cubic equation for the square of the frequency was solved using the Cardan method to find the first root, and the remaining quadratic to find the other two roots. A check on the accuracy of this method of calculating the frequency showed that six figures of accuracy were present for most cases. In the case of two equal roots the accuracy could be reduced to five figures; however, an error of this magnitude is not significant in the construction of the frequency histogram. The calculations were performed on an IBM-7090 computer.

RESULTS

The elastic constants for vanadium at 300°K are given by Alers[12] as

$$c_{11}=22.795\times10^{11}\ \text{dyne-cm}^{-2},$$
$$c_{12}=11.870\times10^{11}\ \text{dyne-cm}^{-2},$$
$$c_{44}=4.255\times10^{11}\ \text{dyne-cm}^{-2}.$$

The maximum frequency calculated from Eq. (11), using these elastic constants, is 8.87×10^{12} cps which compares well with the experimental value of 9×10^{12} cps extrapolated from the curve given by Eisenhauer.[5] The two constant central-force model gives a maximum frequency of 7.07×10^{12} cps. Thus the three constant model gives a better result for the maximum frequency. The calculated and experimental frequency distributions are, however, quite different with regard to the location of the two major peaks, with the calculated frequency distribution having more widely spread maxima. The parameter γ_2/γ_1 may be varied without destroying the good agreement with the experimental maximum frequency. Figure 2 shows the frequency historgrams for vanadium for ratios of γ_2/γ_1 of 0.5 and 1.0, along with the case for $\gamma_2=0$. Figure 2 indicates that increasing the ratio of γ_2/γ_1 reduces the intensity of the first peak and moves the two maxima closer together producing better agreement with experiment. In view of those results, the frequency histogram for $\gamma_2/\gamma_1=1.5$ was also calculated and a smoothed curve obtained from these results is shown in Fig. 3 together with the experimental curve of Eisenhauer. For comparison the smoothed curve for the central-force model of Clark is also presented.

To indicate the effect that an uncertainty regarding the elastic constants might play in the determination

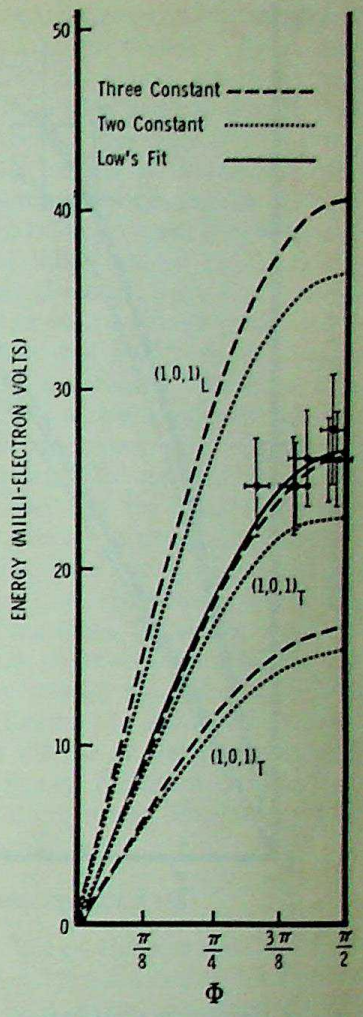


FIG. 4. Dispersion curves for iron along the [101] direction, $\Phi_1=\Phi_3=\Phi$, $\Phi_2=0$ where $\Phi_1=\pi ak_1$, along with the experimental results of Low, and Low's best fit.

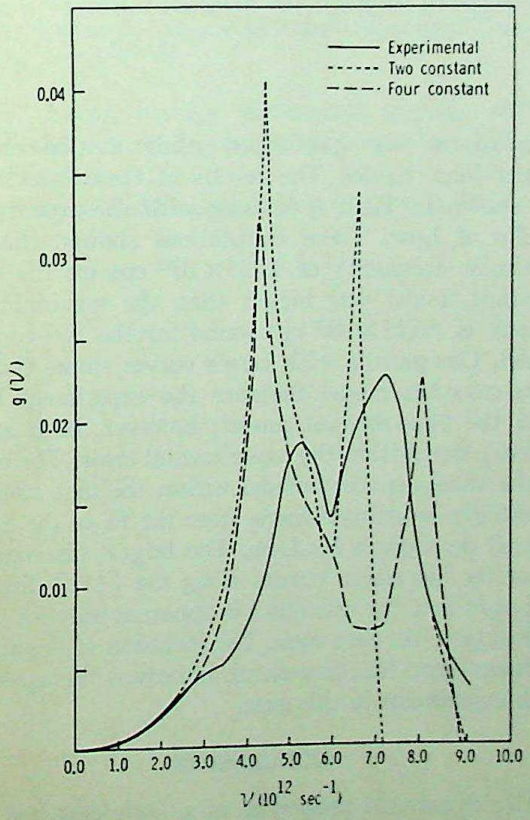


FIG. 3. The experimental results of Eisenhauer presented with the frequency distribution of Clark and the model used here with $\gamma_2/\gamma_1=1.5$. The results are normalized such that the area under the curves are the same.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

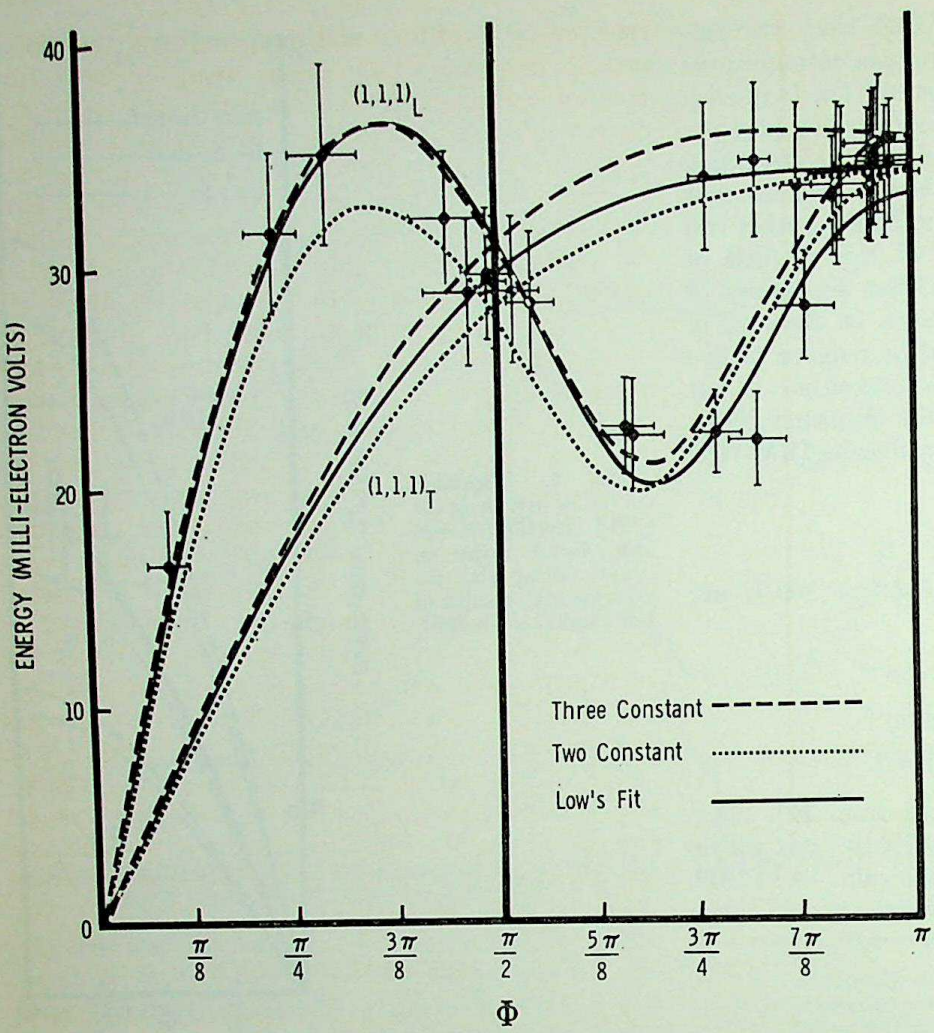


FIG. 5. Dispersion curves for iron along the [111] direction, $\Phi_1=\Phi_2=\Phi_3=\Phi$, where $\Phi_1=\pi a k_1$ along with the experimental results of Low, and Low's best fit.

of the frequency distribution, vanadium distributions were calculated using elastic constants varied by $\pm 10\%$ independently and assuming $\gamma_2=0$. Briefly, it was noted that increasing c_{12} and c_{11} by 10% tends to spread the peaks while increasing c_{44} brings the maxima together slightly. It was also apparent that increasing c_{11} and c_{44} by 10% increased the sharpness of the low-frequency peak while increasing c_{12} had the opposite effect. None of these frequency distributions seem to fit the experimental one appreciably better than the distribution with the unchanged elastic constants, and the calculated maximum frequencies are all in the neighborhood of the experimental maximum.

In addition to the frequency spectra discussed above, the model was used to calculate frequency dispersion curves. The calculations were compared with dispersion curves for iron as determined by Low.[6] Using the elastic constants given by Low for iron at 16°C as

$$c_{11}=2.332\times10^{12}\ \text{dyne-cm}^{-2},$$
$$c_{12}=1.355\times10^{12}\ \text{dyne-cm}^{-2},$$
$$c_{44}=1.180\times10^{12}\ \text{dyne-cm}^{-2},$$

dispersion curves were calculated along the [101], [111], and [001] directions. For comparison the same calculations were performed using a two-constant central-force model. The results of these calculations are shown in Figs. 4–6 along with the experimental results of Low. These calculations showed that the maximum frequency of 9.735×10^{12} cps for the three-constant model was higher than the maximum frequency of 8.711×10^{12} cps found for the two-constant model. Comparison with Low's curves shows that the three-constant model matches the experiment better than the two-constant model; however, both are, in general, well within the experimental error. The results of the three-constant model within the first zone, are essentially indistinguishable from the fit to the experimental data given by Low. The largest discrepancies occur for dispersion curves along the [111] direction, and even here the difference is apparent only for values of Φ outside the first zone. The inclusion of the angular stiffness terms has, in general, improved the agreement with experiment in this case.

CONCLUSION

The dynamical properties of a body-centered cubic lattice of identical particles have been investigated using the Born-von Karman model of a crystal lattice. The model employed contained noncentral as well as

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

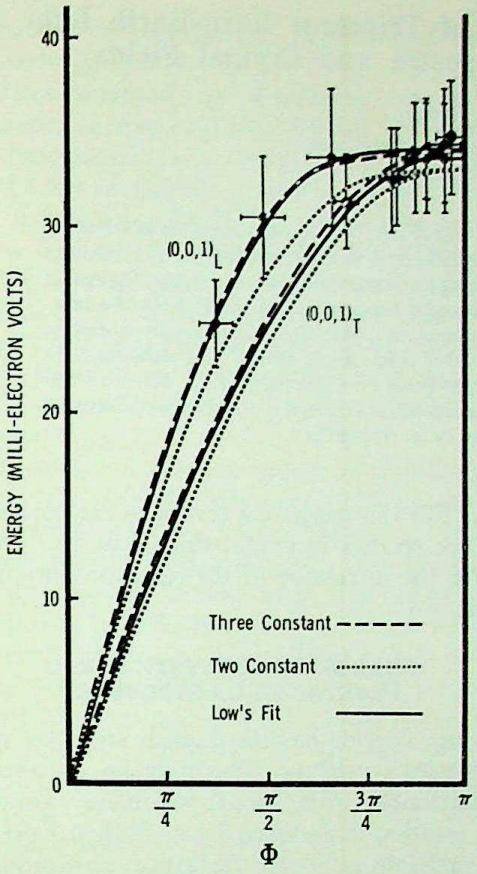


FIG. 6. Dispersion curves for iron along the [001] direction, $\Phi_1=\Phi_2=0$, where $\Phi_3=\pi ak_3$, along with the experimental results of Low, and Low's best fit.

central forces, namely, noncentral angular stiffness forces and central forces between nearest and next-nearest neighbors.

Frequency distribution histograms have been constructed for vanadium and compared with the experimental frequency distributions available from slow neutron experiments. The calculated maximum frequency for vanadium was in good agreement with that derived from the experimental frequency distributions, an improvement over the results of the central-force model. Our model containing a single angular stiffness interaction ($\gamma_2/\gamma_1=0$) gives a frequency distribution which is very similar to that obtained by Hendricks *et al.*[2] with their noncentral model. Both distributions yield a separation between the two major peaks which is significantly larger than that found experimentally. Inclusion of a second angular stiffness interaction into our model ($\gamma_2/\gamma_1\neq 0$) gives a small improvement but not enough to produce agreement with experiment. The disagreement in the experimental and theoretical values of the separation between the two major peaks in the frequency distribution might possibly be removed by including additional interactions in the model. It may well be; however, that short-range forces are inadequate, and that long-range forces are required. Since vanadium is a metal, the effects of the electron gas may be significant as discussed by DeLaunay.[7] The recent calculation of Hendricks, Riser, and Clark,[2] however, indicates that the use of DeLaunay's model does not resolve this difficulty. Failure of the adiabatic approximation in a metal might also be a factor which should be taken into account.

ACKNOWLEDGMENTS

We wish to thank Dr. Robert Herman and Dr. David H. Evans for their many helpful discussions throughout the course of this work. In addition, one of us (B.C.C.) is grateful to Professor Basil Curnutte for his constant encouragement.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Paramagnetic Resonance and Relaxation of Trivalent Rare-Earth Ions in Calcium Fluoride. I. Resonance Spectra and Crystal Fields

M. J. Weber and R. W. Bierig
Raytheon Research Division, Waltham, Massachusetts

(Received 3 January 1964)

The energy level structures arising from the crystal-field splitting of the free-ion ground states of trivalent rare-earth ions in CaF_2 are determined from a survey of the optical and electron paramagnetic resonance spectra. The predicted ground states are in agreement with existing paramagnetic resonance data. Estimates are made of the magnitude of the cubic crystal-field parameters and their variation throughout the rare-earth series. Axial crystal fields of tetragonal and trigonal symmetries, due to nearest-neighbor charge compensation, are found to cause large perturbations on the cubic-field energy levels. The paramagnetic resonance spectra of all rare-earth ions are reviewed and new resonances for Ce^{3+}, Sm^{3+}, Er^{3+}, and Yb^{3+} are reported. Expressions for the calculation of crystal-field matrix elements in terms of 3-j and 6-j symbols and their relationship to the operator equivalent approach are given in an Appendix.

INTRODUCTION

RARE-EARTH ions of the lanthanide series can be introduced into a calcium fluoride host lattice substitutionally at divalent calcium sites. Although several of the rare earths can be stabilized in the divalent state in alkaline earth halide lattices, we shall be concerned almost exclusively with the trivalent rare-earth ions. In the case of the trivalent rare earths, a question arises regarding the compensation of the additional positive charge associated with a trivalent ion substituted in a divalent site. Several different charge compensation mechanisms have been found in CaF_2 which depend upon the growth conditions, impurity concentrations, and thermal treatments of the sample. The crystalline electrostatic fields at the rare-earth sites have correspondingly different magnitudes and symmetries. In this paper and the one to follow[1] (henceforth denoted as II), the electron paramagnetic resonance and spin-lattice relaxation properties of rare-earth ions in sites of different crystal-field symmetries will be examined and compared.

The interpretation of the spin-lattice relaxation results in II requires knowledge of the strength of the crystal fields, the energy level structure of the resonant ion, and the wave functions of the lower states. Unfortunately, such information for the trivalent rare-earth series in CaF_2 is still incomplete. We have made estimates of the cubic crystal-field parameters, their variation throughout the lanthanide series, and the relative size and effect of axial field perturbations from a survey of available optical spectra of rare-earth ions in CaF_2. This information, while not precise, is nevertheless valuable in analyzing paramagnetic resonance spectra, calculating the magnitude of spin-lattice coupling, and estimating the crystal-field splittings when direct measurements are lacking. The predicted crystal-field splitting, ground state, and excited states of the lowest J multiplet of rare earths in various symmetry sites are correlated with the findings of paramagnetic resonance studies. For this purpose a review of the paramagnetic resonance spectra is made which will also provide a basis for the discussion of the relaxation properties of the ions in II.

CHARGE COMPENSATION AND CRYSTAL-FIELD SYMMETRY

Calcium fluoride has the fluoride structure with O_h^5 space group symmetry. This may be visualized as a cubic lattice of fluorine ions in which every other body-center position is occupied by a divalent calcium ion. Rare-earth ions can enter the lattice substitutionally at Ca^{2+} sites. At such sites divalent rare-earth ions experience a cubic crystalline electric field of 8-fold fluorine coordination. Trivalent rare earths, however, experience crystal fields of differing strengths and symmetries depending upon how and where the additional positive charge is compensated. The occurrence of the different charge compensation mechanisms is dependent upon many factors associated with the total crystal growth process. Several resultant distributions of ions which generate crystal-field environments of cubic, tetragonal, trigonal, and orthorhombic symmetry are described

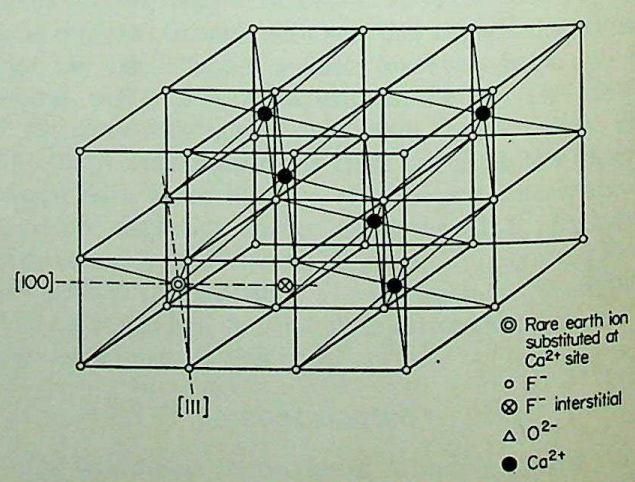


FIG. 1. A portion of the CaF_2 lattice showing the site of substitutional trivalent rare-earth ions and locations of charge compensating ions.

[1] R. W. Bierig, M. J. Weber, and S. I. Warshaw, Phys. Rev. 134, A1504 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

below and can be visualized by inspection of the CaF_2 lattice shown in Fig. 1.

The most commonly found trivalent rare-earth site in CaF_2 is one of tetragonal (C_{4v}) crystal-field symmetry. Charge compensation is attained by the presence of an F^- ion in one of the six nearest-neighbor interstitial sites at a distance $a/2=2.73$ Å, where a is the lattice constant. The three crystal-field symmetry axes are the [100] crystallographic directions. This charge compensating mechanism has been well substantiated from investigations of density and lattice constants by x-ray methods,[2,3] ionic conductivity,[4] as well as by the early paramagnetic resonance results of Bleaney *et al.*[5] and others.[6]

Paramagnetic resonance spectra exhibiting cubic symmetry are observed when charge compensation is not achieved locally. Since the excess F^- in the interstitial site is not tightly bound to the rare earth, it can diffuse through the crystal at high temperatures.[4] Friedman and Low[7] have used this mobility to convert from tetragonal to cubic sites. The sample is heated to disorder the rare-earth $-F^-$ pairs and then rapidly quenched. Whether the interstitial fluorine can diffuse to an energetically favorable site near a trivalent rare-earth ion or gets trapped elsewhere in the lattice depends upon the rate of cooling. Trivalent rare-earth ions of the second half of the $4f^n$ shell have been found in sites of cubic symmetry more frequently than ions of the first half. The stability of the former in cubic sites may be associated with their smaller ionic radii and different local distortions.

A crystal field of trigonal symmetry (C_{3v}) arises from the replacement of one of the eight nearest-neighbor F^- at a distance $\sqrt{3}a/4=2.36$ Å by O^{2-}, a mechanism discussed by Feofilov.[8,9] The symmetry axes are the four [111] cube-body diagonals. The formation of this type of site is a function of the oxidation-reduction conditions during growth. Sierro[10,11] has studied the effects of thermal treatment and hydrolysis at high temperatures on the trigonal resonance spectra of rare earths in alkaline earth halides. If samples are heated in an atmosphere containing a small amount of water vapor, H_2O molecules dissociate at the surface, OH^- replacing F^- which goes off as HF. Trigonal spectra corresponding to OH^- substituted for F^- have been observed. With further thermal treatment, OH^- dissociates, the resulting O^{2-} diffusing through the crystal to compensate a trivalent rare-earth ion as above.

A paramagnetic resonance spectrum having trigonal symmetry has been observed[12,13] from U^{4+} in CaF_2, which is attributed to the replacement of two F^- ions at opposite corners of the fluorine cubic by two O^{2-} ions. Although a resonance from a lanthanide series ion in a similar site has not been reported, Ce, Pr, and Tb are known to exhibit tetravalent states.[14,15]

Recently an orthorhombic spectrum of U^{3+} in CaF_2 has been reported by Mahlab *et al.*[16] The X-Y axes are in a (110) plane, the X axes being at an angle of 19° with respect to [110]. There are twelve inequivalent magnetic sites. Several models for the charge compensation are suggested, the most likely one being an F^- in a fourth nearest-neighbor interstitial site.

In addition to lines which can be explained by the above charge compensation mechanisms, other lines have appeared in both the optical and paramagnetic resonance spectra of rare-earth ions in CaF_2, particularly at higher rare-earth impurity concentrations. From an inspection of Fig. 1, it is not difficult to conjure up other charge compensating ion distributions at sites more distant than described above which may possibly produce measurable deviations from cubic-field symmetry. For example, F^- ions in second nearest-neighbor interstitial sites could produce a [111] trigonal spectrum. [*Note added in proof.* Such trigonal spectra have been reported for Yb^{3+} in SrF_2 and BaF_2 by U. Ranon and A. Yaniv, Phys. Letters 9, 17 (1964).] The replacements of two Ca^{2+} ions by a trivalent rare earth and a monovalent impurity,[17] or three Ca^{2+} ions by two trivalent rare-earth ions have been suggested.[18] The appearance of many new lines as the rare-earth concentration is increased may be due to the presence of energetically favored pairs or clusters of ions. Some of the lines may also be associated with other impurity ions or crystal defects.

Rare-earth-doped, alkaline earth halides exhibit additional interesting properties. For example, several trivalent rare-earth impurities have been converted to divalent ions by exposure to ionizing radiation,[15,19–21] the resulting divalent ion being in a field of cubic sym-

---

[2] E. Zintl and A. Udgard, Z. Anorg. Chem. 240, 150 (1939).
[3] R. W. M. D'Eye and F. S. Martin, J. Chem. Soc. 349, 1847 (1957).
[4] R. W. Ure, J. Chem. Phys. 26, 1363 (1957).
[5] B. Bleaney, P. M. Llewellyn, and D. A. Jones, Proc. Phys. Soc. (London) B69, 858 (1956).
[6] J. M. Baker, W. Hayes, and M. C. M. O'Brien, Proc. Roy. Soc. (London) A254, 273 (1960).
[7] E. Friedman and W. Low, J. Chem. Phys. 33, 1275 (1960).
[8] P. P. Feofilov, Dokl. Akad. Nauk SSSR 99, 731 (1954).
[9] I. V. Stepanov and P. P. Feofilov, Dokl. Akad. Nauk SSSR 108, 615 (1957) [English transl.: Soviet Phys.—Doklady 1, 350 (1957)].
[10] J. Sierro, J. Chem. Phys. 34, 2183 (1961).
[11] J. Sierro, Phys. Letters 4, 178 (1963).
[12] R. S. Title, P. P. Sorokin, M. J. Stevenson, G. D. Pettit, J. E. Scardefield, and J. R. Lanekard, Phys. Rev. 128, 62 (1962).
[13] A. Yariv, Phys. Rev. 128, 1588 (1962).
[14] F. H. Spedding, International Science and Technology, No. 4, 39 (1962).
[15] W. Low and U. Ranon, *Paramagnetic Resonance* (Academic Press Inc., New York, 1963), p. 167.
[16] E. Mahlab, V. Volterra, W. Low, and A. Yariv, Phys. Rev. 131, 920 (1963).
[17] B. Bleaney, J. Appl. Phys. 33, Suppl. I, 358 (1962).
[18] N. Rabbiner, Phys. Rev. 130, 502 (1963).
[19] W. Hayes and J. W. Twidell, J. Chem. Phys. 35, 1521 (1961).
[20] W. Hayes, G. D. Jones, and J. W. Twidell, Proc. Phys. Soc. (London) 81, 371 (1963).
[21] W. Hayes and J. W. Twidell, Proc. Phys. Soc. (London) 82, 330 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

metry. Observations[22] of optical absorption and conductivity due to free electrons in single crystals of CdF_2 doped with rare earths indicate charge compensation by means of electrons in conduction bands or shallow traps. Recently, optical detection of phototransfer of electrons in (Sm, Eu) doped CaF_2 of the form,

$$Eu^{2+}+Sm^{3+}\xrightarrow{h\nu}Eu^{3+}+Sm^{2+},$$

has been reported[23] and interpreted as due to electrons localized at vacancies in the vicinity of the trivalent rare earth. Photoconductivity of divalent rare-earth-doped CaF_2 has also been studied.[24]

CRYSTAL-FIELD PARAMETERS

Information about the crystalline electric field splitting of the free-ion states can usually be derived from studies of optical absorption and fluorescence spectra. For the trivalent rare-earth series in CaF_2, such data is in general incomplete, widely scattered, and a detailed interpretation frequently lacking or not warranted. Among the factors contributing to this status[25,26] are that the spectra are complicated due to the presence of vibrational satellite lines, complex structure appears at moderate-to-high impurity concentrations, different types of spectra exist which are sensitive to changes in neighboring ion distributions and growth conditions, the centroids of crystal-field levels from a given state for different point symmetries do not coincide, the spectral intensities may be weak, and there is no polarized absorption to aid in the identification of cubic-field energy levels.[27,28]

The determination of crystal-field parameters from optical data requires that the site symmetry and energy level assignments be established. Subsequent fitting procedures and numerical computations employed to obtain consistent and meaningful parameters are frequently nontrivial for the multilevel rare-earth energy structures. It may be necessary to include crystal-field admixing between states of different quantum number J and the breakdown of ideal Russell-Saunders coupling. Such detailed calculations are, with a few possible exceptions, not warranted for rare earths in CaF_2, since much of the published data is not complete and the site symmetries are not well established. We shall therefore content ourselves in the next section with less refined estimates of the crystal-field parameters.

The optical spectra of several divalent rare-earth ions in cubic sites in CaF_2 have been studied. Although the crystal-field parameters for these ions will differ from those of their isoelectronic trivalent neighbor, due principally to the influence of the larger effective nuclear charge, they can be of value if a suitable correlation can be found. Since some empirically derived guidance is available for extrapolating the crystal-field parameters of a divalent ion to the corresponding isoelectronic trivalent ion, we shall include several divalent ions in the survey of optical spectra.

Crystal-field parameters can also be obtained from analysis of such factors as the g values, g-value shifts, and hyperfine constants observed in paramagnetic resonance experiments.[29] The results, however, are in general quite sensitive to the accuracy of the measurements; for example, a small uncertainty in g value can give rise to a large variation in the derived crystal-field parameters. We shall be interested principally in a qualitative correlation between paramagnetic resonance data and the crystal-field parameters determined from optical spectra.

The Hamiltonians describing the crystal-field potentials of cubic, tetragonal, and trigonal symmetry may be written as

$$\mathcal{H}_{\text{cubic}}=B_4(O_4^0+5O_4^4)+B_6(O_6^0-21O_6^4), \quad (1)$$

where $B_4^4=5B_4^0$, $B_6^4=-21B_6^0$,

$$\mathcal{H}_{\text{tetra}}=B_2^0O_2^0+B_4^0O_4^0+B_4^4O_4^4+B_6^0O_6^0+O_6^4O_6^4, \quad (2)$$

and

$$\mathcal{H}_{\text{trig}}=B_2^0O_2^0+B_4^0O_4^0+B_4^3O_4^3+B_6^0O_6^0+B_6^3O_6^3+B_6^6O_6^6. \quad (3)$$

The parameters B_n^m are related to the strength of the crystal-field components and, in the notation of Elliott and Stevens,[30] are equal to $\zeta_n A_n^m\langle r^n\rangle$, where ζ_n is the operator equivalent factor, α, β, γ for $n=2$, 4, 6, respectively. The O_n^m's are angular momentum operators transforming as the corresponding spherical harmonics.

In a very useful paper, Lea, Leask, and Wolf[31] (henceforth referred to as LLW) have determined the eigenfunction and eigenvalue solutions of Eq. (1) for various angular momentum J. Their diagrams of the eigenvalues are plotted as a function of a quantity x defined by

$$\frac{x}{1-|x|}=\frac{F(4)B_4}{F(6)B_6}=\frac{F(4)\beta A_4\langle r^4\rangle}{F(6)\gamma A_6\langle r^6\rangle}, \quad (4)$$

---

[22] J. D. Kingsley and J. S. Prener, Phys. Rev. Letters 8, 315 (1962).

[23] P. P. Feofilov, Opt. i Spektroskopiya 12, 531 (1962) [English transl.: Opt. Spectry. (USSR) 12, 296 (1962)].

[24] C. H. Anderson and Z. J. Kiss, Bull. Am. Phys. Soc. 9, 87 (1964).

[25] Feofilov, in a recent survey paper (Ref. 26) on luminescence of rare-earth ions in alkaline earth halides, suggested that a previous assertion may again be *à propos*, namely, that "the variability and diversity of the spectral fine structure exhibited by many phosphors with line spectra led us to the conclusion that a general individual spectra cannot serve as the basis for interpretation."

[26] P. P. Feofilov, Bull. Acad. Sci., USSR 26, 437 (1962).

[27] Z. J. Kiss, J. Chem. Phys. 38, 1476 (1963).

[28] W. Low, *Advances in Quantum Electronics* (Columbia University Press, New York, 1961), p. 138.

[29] W. Low, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1960), Suppl. 2.

[30] R. J. Elliott and K. W. H. Stevens, Proc. Roy. Soc. (London) A218, 553 (1953).

[31] K. R. Lea, M. J. M. Leask, and W. P. Wolf, Phys. Chem. Solids 23, 1381 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where $-1<x<1$ and the F's are multiplicative factors given by LLW. The ordering of states within a J manifold is dependent upon the ratio of the fourth- and sixth-order cubic-field parameters B_4/B_6, and hence from Eq. (4) upon x. The free-ion ground states of the trivalent rare earths and their decomposition in a cubic field are given in Table I. Also listed are estimated x values obtained using approximate cubic-field parameters for CaF_2 and their variation throughout the rare-earth series, which are discussed later. The predicted or experimentally observed cubic field ground states are given in the final column of Table I.

The determination of the parameters $A_n{}^m\langle r^n\rangle$ requires the evaluation of various crystal-field matrix elements. This may be performed easily for the lowest LSJ state using the "operator equivalent" approach of Stevens[32] and existing tables of matrix elements; however, when treating the crystal-field splitting of excited states $L'S'J'$ and matrix elements between different J states, it is convenient to use a more general approach which utilizes the tabulated 3-j and 6-j symbols.[33] This approach and its relationship to that of operator equivalence are developed in the Appendix. A formula is also given, in terms of 6-j symbols, which relates the operator equivalent factors for two different J states. Thus, given x for a state J, it can be used together with Eq. (4) to derive x for other states J', and thereby, from LLW, the expected ordering of levels.

TABLE I. Crystal-field splitting of the free-ion ground state of trivalent rare-earth ions in a cubic field in CaF_2. The x values are obtained using estimated crystal-field parameters in Eq. (4). * denotes ground state based upon estimated x value.

Rare-earth ion	Configuration ground state	Decomposition in cubic field	Estimated x value	Lowest state in CaF_2
Ce^{3+}	$4f^1$, $^2F_{5/2}$	$\Gamma_7+\Gamma_8$	1.0	Γ_8
Pr^{3+}	$4f^2$, $^3H_4$	$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	0.9	Γ_1, Γ_5*
Nd^{3+}	$4f^3$, $^4I_{9/2}$	$\Gamma_6+2\Gamma_8$	−0.7	Γ_8
Pm^{3+}	$4f^4$, $^5I_4$	$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	−0.8	Γ_1
Sm^{3+}	$4f^5$, $^6H_{5/2}$	$\Gamma_7+\Gamma_8$	1.0	Γ_8
Eu^{3+}	$4f^6$, $^7F_0$	Γ_1		Γ_1
Tb^{3+}	$4f^8$, $^7F_6$	$\Gamma_1+\Gamma_2+\Gamma_3+\Gamma_4+2\Gamma_5$	0.9	Γ_2, Γ_3*
Dy^{3+}	$4f^9$, $^6H_{15/2}$	$\Gamma_6+\Gamma_7+3\Gamma_8$	0.6	Γ_8
Ho^{3+}	$4f^{10}$, $^5I_8$	$\Gamma_1+2\Gamma_3+2\Gamma_4+2\Gamma_5$	−0.5	Γ_3*, Γ_5
Er^{3+}	$4f^{11}$, $^4I_{15/2}$	$\Gamma_6+\Gamma_7+3\Gamma_8$	−0.4	Γ_7
Tm^{3+}	$4f^{12}$, $^3H_6$	$\Gamma_1+\Gamma_2+\Gamma_3+\Gamma_4+2\Gamma_5$	0.6	Γ_2*, Γ_3
Yb^{3+}	$4f^{13}$, $^2F_{7/2}$	$\Gamma_6+\Gamma_7+\Gamma_8$	0.7	Γ_7

ELECTRON PARAMAGNETIC RESONANCE AND OPTICAL SPECTRA

A summary of measured paramagnetic resonance g values for trivalent rare-earth ions in various symmetry sites is given in Table II. Additional features of the resonances together with reported optical spectra, estimated crystal-field parameters and energy level structures, and expected resonances spectra for all rare earths are discussed below.

TABLE II. Paramagnetic resonance g values of trivalent rare-earth ions in CaF_2. Numbers in parenthesis following the g values denote references given in the text.

Rare-earth ion	Cubic[a]	Crystal-field symmetry tetragonal	Trigonal
Ce^{3+}	Γ_8: 2.00, 3.1±0.1 (34) Γ_7: 1.297±0.001 (34) (b)	$g_{\parallel}=3.038\pm0.003$ (35) $g_{\perp}=1.396\pm0.002$ (35)	$g_{\parallel}=2.38\pm0.03$ (c) $g_{\perp}<0.1$ (c)
Nd^{3+}	(42) (b)	$g_{\parallel}=4.412\pm0.008$ (5) $g_{\perp}=1.301\pm0.002$ (5)	
Sm^{3+}	See (46)	$g_{\parallel}=0.907\pm0.01$ (46) $g_{\perp}=0.544\pm0.15$ (46)	
Gd^{3+} Tb^{3+}	g:1.991±0.002 (49)	$g_{\parallel}=g_{\perp}=1.992\pm0.001$ (52) $g_{\parallel}=17.768\pm0.020$ (54) $g_{\perp}<0.25$ (54)	$g_{\parallel}=1.992\pm0.001$ (53) $g_{\parallel}=17.28\pm0.01$ (54) $g_{\perp}<0.25$ (54)
Dy^{3+}	Γ_8: 2.63±0.05 (55) 5.48±0.15 (55) 13.7±0.3 (46) Γ_7: 7.47±0.03 (56)	$g_{\parallel}=1.7\pm0.1$ (55) $g_{\perp}=2.82\pm0.05$ (55)	(I) $g_{\parallel}=16\pm1$ (55) $g_{\perp}<1$ (55) (II) $g_{\parallel}=4.93\pm0.05$ (55) $g_{\perp}=1.50\pm0.05$ (55)
Er^{3+}	$g=6.785\pm0.002$ (35)	(I) $g_{\parallel}=7.76\pm0.02$ (35) $g_{\perp}=6.253\pm0.006$ (35) (d) $g_{\parallel}=6.76\pm0.02$ (35) $g_{\perp}=9.11\pm0.01$ (35) (II) $g_{\parallel}=1.746\pm0.002$ (59) $g_{\perp}=9.16\pm0.01$ (59)	(I) $g_{\parallel}=3.30\pm0.01$ (59) $g_{\perp}=8.54\pm0.02$ (59) (II) $g_{\parallel}=2.206\pm0.007$ (59) $g_{\perp}=8.843\pm0.010$ (59) (b)
Yb^{3+}	$g=3.443\pm0.002$ (18) (b)	$g_{\parallel}=2.423\pm0.001$ (c) $g_{\perp}=3.878\pm0.001$ (c)	$g_{\parallel}=1.323\pm0.001$ (65) $g_{\perp}=4.389\pm0.004$ (65)

[a] g values for Γ_8 resonances are for $H\|[100]$.
[b] See discussion in text.
[c] This paper.
[d] From excited state at ~24 cm^{-1}.

[32] K. W. H. Stevens, Proc. Phys. Soc. (London) A65, 209 (1952).
[33] M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, Jr., *The 3-j and 6-j Symbols* (Technology Press, Massachusetts Institute of Technology, Cambridge, 1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Cerium

The cubic crystal-field ground state of Ce^{3+} in CaF_2 should be a Γ_8 quartet. Resonances from an isolated Γ_8 derived from $J=\frac{5}{2}$ would have g values of Λ, $7\Lambda/3$, and $11\Lambda/3$ ($\Lambda=6/7$ for Ce^{3+}) for a magnetic field $H\|[100]$. Dvir and Low[34] have reported broad, anisotropic resonances at 4°K and 1.2 cm with $g=2.00$ and 3.1 ± 0.1, which were interpreted as due to Ce^{3+} in cubic sites. At high fields a line was observed with $g_{\|}=1.297\pm0.001$ and was ascribed to the excited Γ_7 doublet which has a predicted g value of 10/7. Since the line was seen at 4°K, they conjectured that the cubic-field splitting must be very small, not much larger than the Zeeman energy. As we shall see, however, the cubic field splittings of isolated rare-earth ions in CaF_2 are, in general, very much larger. We have examined several $Ce:CaF_2$ samples but have not been able to positively identify any cubic-field resonances.

The resonance of Ce^{3+} in sites of tetragonal symmetry has been reported by Baker *et al.*[35] The observed g values are derivable from a ground-state doublet of the form $\cos\theta|\pm\frac{5}{2}\rangle-\sin\theta|\mp\frac{3}{2}\rangle$, where $\cos\theta\approx0.91$. Using this result and the tetragonal field Hamiltonian Eq. (2) yields

$$\tan2\theta=(2\times5^{1/2}B_4^4)/(B_2^0+20B_4^0)=1.15. \qquad (5)$$

This is very close to the cubic field value, $\tan2\theta=(5)^{1/2}/2=1.12$, obtained by setting $B_2^0=0$ and $B_4^4=5B_4^0$, thereby implying the tetragonal perturbation may be very small. This, however, is in contradiction to results from other rare earths in CaF_2, where axial fields arising from nearest-neighbor charge compensation greatly perturb the cubic-field energy level structure. Also no evidence of a small energy level splitting has been found from the spin-lattice relaxation of Ce^{3+} in tetragonal sites.[1] Thus the above agreement may be coincidental, particularly since any admixing from the $^2F_{7/2}$ excited state, which should be important for $Ce^{3+}:CaF_2$, has been neglected in obtaining this result and since we do not have sufficient information to evaluate B_2^0, B_4^0, and B_4^4 separately. The above ground-state wave function and the eigenvalues for $J=\frac{5}{2}$ in a tetragonal field, Eq. (6), are consistent with a case where the second-order crystal-field term does not dominate and $A_2^0\langle r^2\rangle$ is positive.

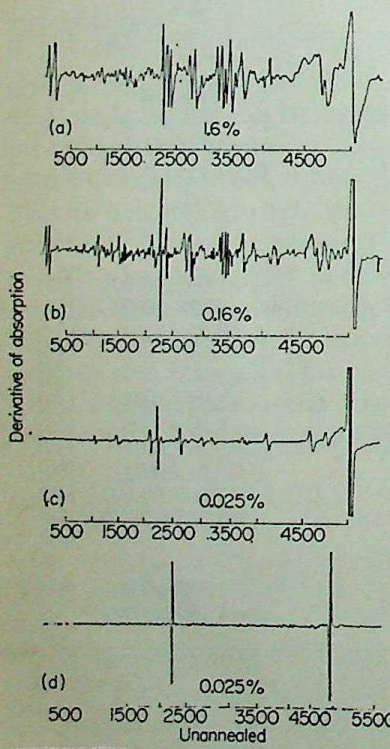


FIG. 2. Paramagnetic resonance spectrum of $Ce^{3+}:CaF_2$ at 9.6 kMc/sec and 4.2°K for various Ce concentrations. Spectrum (d) is from an unannealed sample (see text). The magnetic field, in Oe, is approximately $\|[100]$.

We have observed resonances arising from Ce^{3+} in sites of trigonal symmetry. The g values given in Table II can be compared with those expected for a doublet of the form $|\pm\frac{3}{2}\rangle$ for which $g_{\|}=18/7=2.57$, $g_{\perp}=0$; the differences are probably due mainly to admixing with states from $^2F_{7/2}$. The trigonal field eigenvalues for an isolated $J=\frac{5}{2}$ state are given in Eq. (7).

At higher rare-earth impurity concentrations, many additional lines frequently appear in the paramagnetic resonance spectrum as in the case of the optical spectra.[27] The increasing spectral complexity with concentration is illustrated in Fig. 2 where three spectra from samples[36] containing cerium dopings, each differing by approximately one order of magnitude, are shown. Such additional spectra may be attributed speculatively to the simultaneous existence of several different charge compensation mechanisms and ion arrangements, crystal imperfections, or clusters of ions, but additional work is needed to clarify the situation. The effects of annealing conditions upon the number of lines, linewidth, and line shape is also evident. Traces (a), (b), (c) are from samples which were annealed at 1200°C for 12 h and cooled at a rate of 20°C/h. Trace (d) is from an "as-grown" sample cooled at a rate of ~200°C/h. Note the presence of the two extremely low-field resonances in Figs. 2 (a), (b), which correspond to the zero-field resonances at 9.37 and 9.48 kMc/sec reported by Kaplan.[37]

The optical spectra of $Ce^{3+}:CaF_2$ has been reported by several authors[34,38–40]; however, no complete crystal-field analysis has been attempted. Whereas the free-ion $^2F_{5/2}$ and $^2F_{7/2}$ states are separated by 2250 cm^{-1}, the infrared absorption spectrum[39] of $Ce^{3+}:CaF_2$ is characterized by two broad bands centered around

[34] M. Dvir and W. Low, Proc. Phys. Soc. (London) **75**, 136 (1960).

[35] J. M. Baker, W. Hayes, and D. A. Jones, Proc. Phys. Soc. (London) **73**, 942 (1959).

[36] See Ref. 1 for more information on sample preparation.

[37] D. E. Kaplan, Bull. Am. Phys. Soc. 8, 468 (1963).

[38] P. P. Feofilov, Opt. i. Spektroskopiya **6**, 234 (1959) [English transl.: Opt. Spectry. (USSR) **6**, 150 (1959)].

[39] Z. J. Kiss, Phys. Rev. **127**, 718 (1962).

[40] A. A. Kaplyanskii, V. N. Medvedev, and P. P. Feofilov, Opt. i Spektroskopiya **14**, 664 (1963) [English transl.: Opt. Spectry. (USSR) **14**, 351 (1963)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2200 and 3200 cm^{-1}. Estimates of the fourth- and sixth-order cubic-field parameters for Ce^{3+}, based upon extrapolations of results for other rare-earth ions in CaF_2, suggest an x value for $^2F_{7/2}$ which in LLW is near the crossing of the Γ_6 and Γ_8 energy levels. Assuming the cubic field is dominant, the above lines could correspond to transitions from the $^2F_{5/2}$ Γ_8 ground state to the Γ_6-Γ_8 and Γ_7 states of $^2F_{7/2}$, respectively. At low temperatures several observers have noted a complex line structure about the 2200-cm^{-1} region with a spread $\sim$350 cm^{-1}. In addition to vibrational satellites, this structure could arise from a small Γ_6-Γ_8 separation or a combination of this and additional splitting of the Γ_8 quartet of $^2F_{7/2}$ due to axial crystal fields, since the site symmetries in the samples are not reported.

If one assumes the $\sim$1000 cm^{-1} spectral separation is due mainly to the cubic-field splitting of $^2F_{7/2}$ and equates it to the Γ_7-Γ_6 energy difference and furthermore assumes $x\approx 0.8$, which is near the Γ_6-Γ_8 crossover, crystal-field parameters $A_4\langle r^4\rangle \sim -320$ cm^{-1} and $A_6\langle r^6\rangle \sim 45$ cm^{-1} are obtained. Using these parameters, an energy separation of $\sim$700 cm^{-1} is predicted between the Γ_8 and Γ_7 states of $^2F_{5/2}$. The above results can only be considered as coarse estimates, however, and may differ greatly from the actual values because, assuming the transitions are assigned correctly: (1) the over-all $^2F_{7/2}$ splitting may not be due solely to the cubic field; (2) the x value used was only an estimate; and (3) no account has been made of the off-diagonal elements of the crystal field connecting the $^2F_{5/2}$ and $^2F_{7/2}$ states. The last is certainly not negligible if the crystal field is as large as estimated above. When the site symmetry and various transitions are established, the crystal-field parameters may be determined by solving the combined spin-orbit and crystal-field problem using a representation in which L_z and S_z are diagonal[41] rather than using the more usual perturbation approach with a $|J,J_z\rangle$ basis.

Praseodymium

The cubic-field ground state of Pr^{3+} in CaF_2 should be either a Γ_1 or a Γ_5 state depending upon the ratio of the fourth- and sixth-order crystal-field parameters. The estimated x value in Table I predicts that the magnetic Γ_5 triplet is lowest. The cubic field wave functions describing the Γ_5 state for $J=4$ are

$$(\tfrac{7}{8})^{1/2}|\pm 3\rangle - (\tfrac{1}{8})^{1/2}|\mp 1\rangle,$$
$$(\tfrac{1}{2})^{1/2}(|2\rangle - |-2\rangle).$$

This state behaves like a spin-1 triplet in a magnetic field yielding an isotropic g value of 2.00.

We have searched for resonances at 9.6 kMc/sec and 2–4°K in a CaF_2 sample containing 0.57% Pr. A weak signal was observed with an isotropic g value $\sim$2 which consisted of a single line of width $\sim$20 G superimposed upon a broad line of width $\sim$100 G. Hyperfine structure due to the one stable isotope Pr^{147}, nuclear spin $I=\frac{5}{2}$, which should be a distinguishing feature was not evident. It was not known whether Pr^{3+} was actually present in sites of cubic symmetry in the sample studied. The absence of a cubic resonance could be due to the difficulty encountered in getting trivalent rare earths belonging to the first half of the $4f$ shell into cubic sites. An additional broad intense resonance corresponding to a zero-field splitting of $\sim$0.3 cm^{-1} was also observed which is being investigated further.

Neodymium

The cubic-field ground state of Nd^{3+} in CaF_2 is a Γ_8 quartet for all x ($x<0$). The g values for the Γ_8 resonances depend upon the ratio B_4/B_6 and possible admixing from excited states. A cubic field resonance spectrum for Nd^{3+}:CaF_2 was reported by Vincow and Low[42]; however, there has been some controversy[43] regarding the interpretation of the spectrum which indicates $x\simeq -1.0$. In a tetragonal field the $^4I_{9/2}$ free-ion ground state is decomposed into five Kramers doublets. Resonances for Nd^{3+} in sites of tetragonal field symmetry were found by Bleaney *et al.*[5] and fitted to a wave function of the form $a|\pm\frac{9}{2}\rangle + b|\pm\frac{1}{2}\rangle + c|\mp\frac{7}{2}\rangle$ with $c\approx 0$.

The optical absorption and fluorescence spectra of Nd^{3+} in cubic sites have been studied by Kiss.[27] The crystal-field parameters determined from the measured splittings of the $^4I_{9/2}$ and $^4I_{11/2}$ states, neglecting certain off-diagonal terms, are

$$A_4\langle r^4\rangle = -415 \text{ cm}^{-1}, \quad A_6\langle r^6\rangle = 33 \text{ cm}^{-1}.$$

Using these results in Eq. (4) yields x values of ≈ -0.70 and ≈ -0.91 for $^4I_{9/2}$ and $^4I_{11/2}$, respectively. The observed energy level spacings are similar but do not correspond exactly to those in the diagrams of LLW for these x values, thereby indicating that consideration of the admixing of states of different J may be required to determine satisfactory crystal-field parameters. The cubic field energy levels of $^4I_{9/2}$ are $\Gamma_8^{(2)}-0$, $\Gamma_8^{(1)}-180$ cm^{-1}, and Γ_6-700 cm^{-1}.

We have examined the fluorescence spectrum[44] in a CaF_2 sample which, from an examination of the paramagnetic resonance spectrum, had Nd^{3+} predominantly in tetragonal sites. Transitions between $^4F_{3/2}$ and $^4I_{9/2}$ were observed at 8680, 8750, 8940, 9048, 9078, 9220, and 9465 Å at liquid helium temperatures (these may include vibrational satellites). Lines were also observed at 8860, 9150, and 9335 Å at room temperature which became much smaller or had disappeared at 77°K. In an axial field, the upper fluorescence state $^4F_{3/2}$ decomposes into two Kramers doublets separated by $6B_2^0$, where $B_2^0=\alpha(4f^3, {}^4F_{3/2})A_2^0\langle r^2\rangle$. If the disappearing lines

[41] M. J. M. Leask, R. Orbach, M. J. D. Powell, and W. P. Wolf, Proc. Roy. Soc. (London) A272, 371 (1963).

[42] G. Vincow and W. Low, Phys. Rev. 122, 1390 (1960).
[43] W. P. Wolf, J. Phys. Soc. Japan 17, Suppl. B-I, 442 (1962).
[44] We thank B. DiBartolo for making these measurements.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

are caused by thermal depopulation of the upper doublet, this implies a relatively large tetragonal field splitting and an estimated $|A_2^0\langle r^2\rangle|$ of several hundred cm^{-1}. Such a field would greatly perturb the two Γ_8 quartets of $^4I_{9/2}$. Measurements[1] of the spin-lattice relaxation of Nd^{3+} in tetragonal sites indicate an excited state at $\sim 65\ cm^{-1}$.

Promenthium

Neither the paramagnetic resonance nor optical spectrum of $Pm^{3+}:CaF_2$, which does not occur naturally, has been reported. In a cubic field the lowest state from $^5I_4$ is a Γ_1 singlet for all values of x ($x<0$). The remaining excited states should be relatively far removed.

Samarium

We have searched for resonances at 16 kMc/sec and 4.2°K in an $\sim 0.1\%$ Sm-doped CaF_2 crystal. The sample was prepared so as to enhance the occurrence of trivalent rather than nonresonant divalent samarium ions.[45] The only resonance observed was one exhibiting tetragonal symmetry about the [100] axes which probably arises from interstitial fluorine charge compensation. The measured g values are $g_{11}=0.93\pm0.04$, $g_1=0.57\pm0.05$, which can be fitted by an approximate ground doublet wave function $\cos\theta|\pm\frac{5}{2}\rangle-\sin\theta|\mp\frac{3}{2}\rangle$, where $\cos\theta\approx 0.88$. The hyperfine structure of Sm^{147} and Sm^{149} (nuclear spins $I=\frac{7}{2}$) was observed for $H\|[100]$ where $A^{147}\approx 170\times10^{-4}\ cm^{-1}$. Low[46] has recently confirmed these measurements and has also reported seeing additional resonances exhibiting cubic and trigonal symmetry; however, their interpretation is not yet complete. In the cubic field of CaF_2 the Sm^{3+} ground state is a Γ_8 quartet.

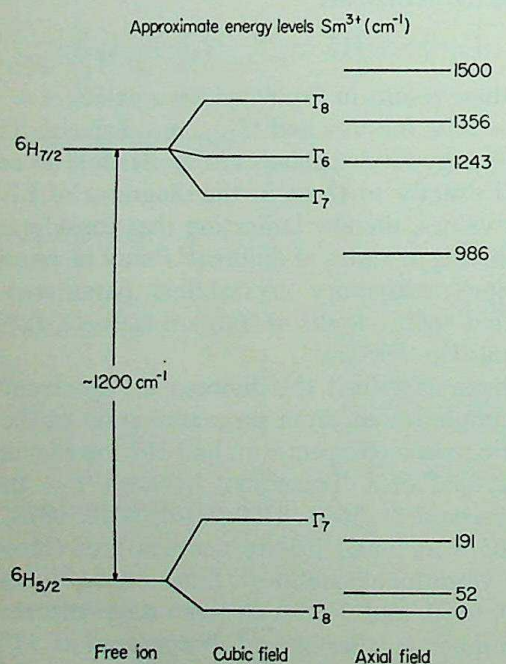


FIG. 3. Crystal-field splitting of the two lowest states of Sm^{3+} in CaF_2. The cubic-field energy level structure is estimated; the axial field values are from Rabbiner (Ref. 18).

A complex fluorescence spectrum from Sm^{3+} in a field of axial symmetry in CaF_2 has been observed and analyzed by Rabbiner.[18] The energy levels for $^6H_{5/2}$ and $^6H_{7/2}$ are shown in Fig. 3. Although no optical spectra of Sm^{3+} in cubic sites have been reported, the ordering of levels within the two lowest ^6H states is expected to be approximately as shown (no significance should be attached to the relative energy level spacings, however, which require knowledge of the cubic field parameters and treatment of crystal-field admixing between J states). From Fig. 3 it can be seen qualitatively that the axial field removes the degeneracy of the Γ_8 states and produces a large perturbation on the cubic-field energy level structure. This is also evident from the resonance results where for tetragonal sites $\tan 2\theta$ in Eq. (5) equals 1.46 compared to the cubic-field value of 1.12.

The symmetry of the axial field present in the above fluorescence spectrum was not established. In a tetragonal field, the energies of the three Kramers doublets for an isolated $J=\frac{5}{2}$ state are, from Eq. (2),

$$\begin{aligned} E(\pm\tfrac{1}{2}) &= -8B_2^0+120B_4^0, \\ E(\pm\tfrac{5}{2},\mp\tfrac{3}{2}) &= 4B_2^0-20B_4^0 \\ &\quad \pm 6[(B_2^0+20B_4^0)^2+20(B_4^4)^2]^{1/2}. \end{aligned} \tag{6}$$

In a trigonal field the three doublets from Eq. (3) are

$$\begin{aligned} E(\pm\tfrac{3}{2}) &= -2B_2^0-180B_4^0, \\ E(\pm\tfrac{5}{2},\mp\tfrac{1}{2}) &= B_2^0+90B_4^0 \\ &\quad \pm 3[(3B_2^0-10B_4^0)^2+10(B_4^3)^2]^{1/2}. \end{aligned} \tag{7}$$

These energies are for isolated J states. The crystal-field splitting for $Sm^{3+}:CaF_2$ in Fig. 3, however, are not negligible compared to the separations of the various J states. Crystal-field admixing must be taken into account in this case to obtain meaningful parameters B_n^m and predicted g values.

The fluorescence spectrum of Sm^{3+} in a tetragonal field of CdF_2 has also been reported.[47] Since CdF_2 has (1) an isomorphic lattice structure, (2) a lattice constant very nearly equal to that of CaF_2, 5.40 Å versus 5.451 Å, and (3) the Pauling ionic radii of Cd^{2+} and Ca^{2+} are approximately equal, 0.97 Å versus 0.99 Å, the crystal-field splittings and ordering of states in the two crystals should be similar. This is indeed found for $Sm^{3+}:CdF_2$, where, in place of the energy levels in Fig. 3, the corresponding levels in CdF_2 from $^6H_{5/2}$ are at 0, 54, and 147 cm^{-1} and from $^6H_{7/2}$ at 959, 1169, 1381, and 1586 cm^{-1}. We shall include data from CdF_2 when similar results for CaF_2 are lacking.

[45] J. R. O'Connor and H. A. Bostick, J. Appl. Phys. 33, 1868 (1962).
[46] W. Low, Phys. Rev. (to be published).
[47] J. S. Prener and J. D. Kingsley, J. Chem. Phys. 38, 667 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Europium

The ground state of Eu^{2+} is a singlet $^7F_0$. From the fluorescence spectrum of Eu^{3+} in isomorphic CdF_2, energy levels from the first excited J state $^7F_1$ are reported[48] at 307–309 and 400 cm^{-1}. In a crystal field of axial symmetry a $J=1$ state is split into a singlet and a doublet with an energy separation $3\alpha A_2{}^0\langle r^2\rangle$. Setting this equal to the observed 92 cm^{-1} separation and using $\alpha(4f^6, {}^7F_1)=-\frac{1}{5}$, yields $A_2{}^0\langle r^2\rangle\approx 150$ cm^{-1}.

Gadolinium

Resonances from Gd^{3+} have been observed in sites of cubic,[49,50] tetragonal,[51,52] and trigonal[10,11,53] symmetry. Since it is an S-state ion, the over-all crystal-field splittings are small, being 0.149 cm^{-1} in a cubic field. In an axial field the $^8S_{7/2}$ ground state is split into 4 doublets with a total splitting of about 2 cm^{-1} for tetragonal and trigonal (O^{2-}) sites.

Terbium

The cubic-field ground state of $Tb^{3+}(^7F_6)$ should be either a Γ_2 singlet or a nonmagnetic Γ_3 doublet, thus no cubic-field resonances are expected nor have they been reported. For $x\gtrsim 0.8$, as estimated in Table I, the ground state is a Γ_3 with a relatively near Γ_5 excited state. Terbium resonances and hyperfine structure for ions in sites of tetragonal and trigonal symmetry have been investigated in detail by Forrester and Hempstead.[54] The resonances for this non-Kramers ion arise from states of principally $|\pm 6\rangle$ character which have a g_{11} of $12\Lambda=18$.

Dysprosium

We have reported[55] the resonance spectrum of Dy^{3+} in sites of cubic, tetragonal, and trigonal symmetry elsewhere. The cubic-field ground state is a Γ_8 quartet with g values corresponding closely to those calculated from the $J=15/2$ wave functions of LLW for $\Gamma_8{}^{(1)}$ and $x=0.6$. The first excited state is a Γ_7 doublet located ~ 8 cm^{-1} above the ground state. Using this energy separation and the above x value, the cubic-field parameters are $A_4\langle r^4\rangle=-(200\text{–}300)$ cm^{-1} and $A_6\langle r^6\rangle=30\text{–}40$ cm^{-1}. The large uncertainty in the values arises from the fact that since we are near the crossover of the Γ_7 and Γ_8 levels and are using the Γ_7-Γ_8 energy separation as a scale factor, a small change in x can give rise to a large change in the over-all crystal-field splitting of the ground state. Recently, Low[56] has repeated measurements of the cubic-field spectrum at higher frequency and fields with accordingly more accuracy. He finds $A_4\langle r^4\rangle=-242\pm 20$ cm^{-1}, $A_6\langle r^6\rangle=41\pm 5$ cm^{-1}, where any admixing from $^6H_{13/2}$ at ~ 3500 cm^{-1} was neglected.

Rabbiner[57] has reported measurements of the fluorescence spectrum of Dy^{3+}:CaF_2. The assigned cubic-field splitting of $^6H_{15/2}$ is ≈ 615 cm^{-1} with a Γ_7 ground state and a first excited Γ_8 state at 53 cm^{-1}. The discrepancy with the above may arise from the presence of additional spectral lines from ions in different symmetry sites.

Holmium

The possible cubic-field ground states of Ho^{3+} in CaF_2 are a nonmagnetic $\Gamma_3{}^{(2)}$ doublet and a $\Gamma_5{}^{(2)}$ triplet (g value ≈ 5). In the region $0.5\lesssim x$, where the Γ_3 state is lowest, the first excited state is a $\Gamma_4{}^{(2)}$ triplet which in the energy level diagram of LLW is nearly coincident with the $\Gamma_3{}^{(2)}$ doublet.

We have searched for a resonance at X band and 2–4°K in a CaF_2 sample grown by Optovac, Inc., containing a nominal Ho concentration of $\sim 0.1\%$. No resonances were observed. Hayes *et al.*[20] have searched for resonances at K band and 4°K in 0.01 and 1% Ho-doped CaF_2 samples which exhibited optical absorption spectra charactersitic of Ho^{3+}. Again no resonances were found suggesting the ground state is a Γ_3 state.

Dy^{2+} is isoelectronic with Ho^{3+}. At 77°K the fluorescence spectrum[58] of Dy^{2+} from $^5I_7$ to the $^5I_8$ ground state exhibits a group of lines between 0–100 cm^{-1} and another group around 400–500 cm^{-1}. Laser action has been observed to a terminal state 90 cm^{-1} above the ground state. From an examination of the energy level diagram for $J=8$ in LLW, the above spectrum indicates that $-0.6<x<-0.4$. Assuming an over-all crystal-field splitting of 450 cm^{-1} for $^5I_8$ and an x value of -0.5, yields cubic-field parameters of $A_4\langle r^4\rangle\sim -400$ cm^{-1} and $A_6\langle r^6\rangle\sim 40$ cm^{-1} for Dy^{2+}.

Erbium

Resonances from Er^{3+} in sites of cubic,[35] tetragonal (I)[35,59] and (II),[59] and trigonal (I) and (II)[59] symmetry have been reported. We have examined several CaF_2 crystals at X-band frequencies containing 0.1–1.0% Er and have identified cubic, tetragonal (I) and trigonal (II) resonances. The spectrum at these frequencies and moderately high Er concentrations is complex and consists of many broad, anisotropic lines with overlapping

[48] J. D. Kingsley and J. S. Prener, Phys. Rev. **126**, 458 (1962).
[49] W. Low, Phys. Rev. **109**, 265 (1958).
[50] C. Ryter, Helv. Phys. Acta **30**, 353 (1957).
[51] J. Sierro and R. Lacroix, Compt. Rend. **250**, 2686 (1960).
[52] V. M. Vinokurov, M. M. Zaripov, Yu. T. Pol'skii, V. G. Stepanov, G. K. Chirkin, and L. Ya. Shekun, Fiz. Tverd. Tela **4**, 2238 (1962) [English transl.: Soviet Phys.—Solid State **4**, 1637 (1963)].
[53] V. M. Vinokurov, M. M. Zaripov, Yu. E. Pol'skii, V. G. Stepanov, G. K. Chirkin, and L. Ya. Shekun, Fiz. Tverd. Tela **5**, 599 (1963) [English transl.: Soviet Phys.—Solid State **5**, 436 (1963)].
[54] P. A. Forrester and C. F. Hempstead, Phys. Rev. **126**, 923 (1962).
[55] R. W. Bierig and M. J. Weber, Phys. Rev. **132**, 164 (1963).
[56] W. Low, Proc. Phys. Soc. (London) **76**, 307 (1960).
[57] N. Rabbiner, Phys. Rev. **132**, 224 (1963); **133**, 1 (1964).
[58] Z. J. Kiss and R. C. Duncan, Proc. IRE **50**, 1531 (1962).
[59] U. Ranon and W. Low, Phys. Rev. **132**, 1609 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

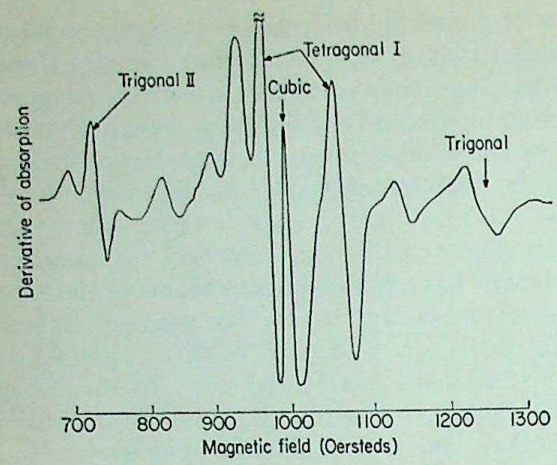


FIG. 4. Low-field portion of the paramagnetic resonance spectrum of Er^{3+}: CaF_2 arising from ions in cubic, tetragonal, and trigonal crystal fields. Hyperfine structure from Er^{167} ($I=\frac{7}{2}$) is also present. Measurements were made at 9.58 kMc/sec, 4.2°K, and with $H\|[110]$.

hyperfine structure, as evident from the low-field portion of the spectrum shown in Fig. 4. Studies of the angular dependence of the spectrum in both the (100) and (111) planes reveal the presence of several unaccounted for resonances which do not correspond to those reported above. One resonance exhibits trigonal symmetry with $g_{\|}=6.31\pm0.05$ and $g_{\perp}=2.14\pm0.04$; however, these g values do not satisfy the relationship $(g_{\|}+2g_{\perp})\approx 3g_{\text{cubic}}(\Gamma_7)$ as do the axial symmetry Er^{3+} resonances observed previously.[59] The total spectrum was examined with increasing temperature up to ~50°K. The trigonal (II) resonance exhibited rapid line broadening in the vicinity of 30°K. The cubic-tetragonal (I) group of lines also exhibited broadening, the cubic line showing the more rapid temperature dependence. By 40°K the cubic and $g_{\|}$, $g_{\perp}$ tetragonal resonances could no longer be distinguished as separate lines. No definite evidence of the tetragonal (I) excited-state resonance reported by Baker *et al.*[35] was noted, although it may have been partially masked by other lines in the spectrum. No other excited-state resonances were detected. Measurements[1] of the temperature dependence of T_1 for the trigonal (II) resonance indicate the presence of an excited state at ~ 50 cm^{-1}.

Paramagnetic resonance studies[60] have shown that the cubic-field ground state of Ho^{2+} in CaF_2 is a Γ_6 doublet in contrast to the Γ_7 ground state found for isoelectronic Er^{3+}. Since the change in the ratio of the fourth- and sixth-order cubic-field parameters is expected to be small between a divalent ion and its isoelectronic trivalent neighbor, the above result suggests that for both Ho^{2+} and Er^{3+} we are in the vicinity of the crossover of the Γ_6 and Γ_7 levels which for $J=15/2$ occurs at $x=-0.46$.

The lowest excited crystal-field states in Ho^{2+}: CaF_2 are[61] a quartet at 30.1 cm^{-1} and a doublet at 33.8 cm^{-1} which from LLW suggest an x value ≈ -0.52. Using these values, the crystal-field parameters for Ho^{2+} are $A_4\langle r^4\rangle\approx -190$ cm^{-1} and $A_6\langle r^6\rangle\approx 16$ cm^{-1}, where any admixture from $^4I_{13/2}$ is probably small and has been neglected. Since x is smaller for Er^{3+}, the sixth-order term must be correspondingly more important.

Pollack[62] has reported a total separation of about 400 cm^{-1} for the 8 crystal-field levels of $^4I_{15/2}$ for Er in an axial field. As an order-of-magnitude estimate, assume a total cubic-field splitting of approximately the same size, 400 cm^{-1}, and an x value of -0.4. These yield $A_4\langle r^4\rangle\approx -310$ cm^{-1} and $A_6\langle r^6\rangle\approx 44$ cm^{-1}, which, while not based upon direct spectroscopic data, are, nevertheless, consistent with expected values extrapolated from results for other ions.

Thulium

The cubic-field ground state of Tm^{3+} may be either Γ_2 or Γ_3. The estimated x value in Table I predicts a Γ_2 singlet lowest. Hayes and Twidell[19] did not observe any resonances in a 0.05% Tm sample at X band and 20°K, but did observe a Tm^{2+} resonance after x irradiation. It has been reported that an axial spectrum of Tm^{3+} has been seen.[15,63]

Ytterbium

The ground state of Yb^{3+} in cubic sites in CaF_2 is a Γ_7 doublet. The wave function describing the Γ_7 doublet for $J=\frac{7}{2}$ is $(\sqrt{3}/2)|\pm\frac{5}{2}\rangle-(\frac{1}{2})|\mp\frac{3}{2}\rangle$ which yields an isotropic g-value of $24/7=3.429$. Low[64] has measured $g=3.426\pm0.001$; Hayes and Twidell[19] report 3.443 ± 0.002. We find a g value of 3.441 ± 0.004. This resonance is observable up to temperatures ~100°K, which implied a comparatively long spin-lattice relaxation time.[1]

We have also observed resonances due to Yb^{3+} in tetragonal sites. The g values in Table II are consistent with a doublet derived from the cubic Γ_7, since $(g_{\|}+2g_{\perp})\approx 3g_{\text{cubic}}(\Gamma_7)$ and indicate that the other states of $^2F_{7/2}$ are sufficiently far removed to have little effect on the g values.[60] They can be fitted by an approximate doublet wave function $0.806|\pm\frac{5}{2}\rangle-0.591|\mp\frac{3}{2}\rangle$. Two trigonal field resonances have been reported by Low and co-workers.[21,65]

Kiss[39] has investigated the absorption and fluorescence spectrum of isoelectric Tm^{2+} in CaF_2. The $^2F_{7/2}$ ground state is split into a Γ_7 doublet lowest, a Γ_6 doublet at 547 cm^{-1}, and a Γ_8 quartet at 556 cm^{-1}. The cubic-field parameters derived from these values are $A_4\langle r^4\rangle=-180$ cm^{-1} and $A_6\langle r^6\rangle=35$ cm^{-1}, which correspond to an x value of 0.74.

The absorption spectrum of Yb^{3+} was also examined

[60] H. R. Lewis and E. S. Sabisky, Phys. Rev. 130, 1370 (1963).
[61] H. Weakliem and Z. J. Kiss (to be published).
[62] S. A. Pollack, Bull. Am. Phys. Soc. 8, 614 (1963).
[63] W. Low, J. Phys. Soc. Japan 17, Suppl. B-I, 440 (1962).
[64] W. Low, Phys. Rev. 118, 1608 (1960); J. Chem. Phys. 37, 30 (1962).
[65] W. Low and U. Rosenberger, Compt. Rend. 254, 1771 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

briefly by Kiss. Two dominant lines were observed at 10 240 cm^{-1} and 10 870 cm^{-1}, which have also been observed by others.[28,64,66] If these lines are ascribed[39] to transitions to the two cubic-field levels of the $^2F_{5/2}$ excited state, an $A_4\langle r^4\rangle$ value of -280 cm^{-1} is found. Comparing the crystal-field splitting of isoelectronic Tm^{2+} and Yb^{3+}, the larger value for the latter reflects the effect of the additional nuclear charge of the trivalent ion. Approximate cubic-field energy levels for Yb^{3+}, based upon Kiss' assignment and assuming $x\sim0.7$, are shown in Fig. 5 together with expressions from the eigenvalues. The estimated over-all crystal-field splitting of $^2F_{7/2}$ is ~800 cm^{-1}.

Low has investigated the electron paramagnetic resonance spectra[64] and the fine structure in optical spectra[28,64] of Yb:CaF_2 and suggested that the cubic field was relatively weak with an over-all ground-state splitting of ≈12 cm^{-1}. Interpretations based upon such fine structure are, in general, inconsistent with the present results and estimates of the magnitude of the cubic field for other rare-earth ions. The existence of such small splittings for Yb:CaF_2 are also not evident from measurements of the g value[19] and the rate and temperature dependence of the spin-lattice relaxation.[1] Measurements[67] of the g shift, temperature dependence, intensity, and infrared aborption of Yb^{3+} in cubic sites in CdF_2 indicate an excited state at ≈50–70 cm^{-1} and an over-all $^2F_{7/2}$ splitting of ≈116 cm^{-1}. We do not understand why these values are so much smaller than those estimated for Yb:CaF_2. Further work to confirm the measurements and verify the interpretations in the two cases are needed.

The eigenvalues for a $J=\frac{7}{2}$ state in a tetragonal field are given by

$$E(\pm\tfrac{5}{2},\mp\tfrac{3}{2})=-b_2^0-8b_4^0+2b_6^0 \pm[(2b_2^0-5b_4^0-7b_6^0)^2+3(5b_4^4-7b_6^4)^2]^{1/2}, \quad (8)$$

$$E(\pm\tfrac{7}{2},\mp\tfrac{1}{2})=b_2^0+8b_4^0-2b_6^0 \pm[(6b_2^0-b_4^0+3b_6^0)^2+35(b_4^4+3b_6^4)^2]^{1/2},$$

where

$$b_2^0=3B_2^0, \quad b_6^0=1260B_6^0,$$
$$b_4^0=60B_4^0, \quad b_6^4=60B_6^4.$$
$$b_4^4=12B_4^4,$$

The observed ground state is of the form $\cos\theta|\pm\frac{5}{2}\rangle -\sin\theta|\mp\frac{3}{2}\rangle$ for which, from Eq. (2) and the doublet wave function found experimentally,

$$\tan2\theta=\sqrt{3}(5b_4^4-7b_6^4)/(2b_2^0-5b_4^0-7b_6^0)=-3.18. \quad (9)$$

The difference between this value and the cubic-field value $\tan2\theta=-\sqrt{3}$, corresponding to $b_2^0=0$, $b_4^4=b_4^0$, and $b_6^4=-b_6^0$, again reveals qualitatively that nearest-neighbor charge compensation produces a large change in the crystal field at the trivalent rare-earth site.

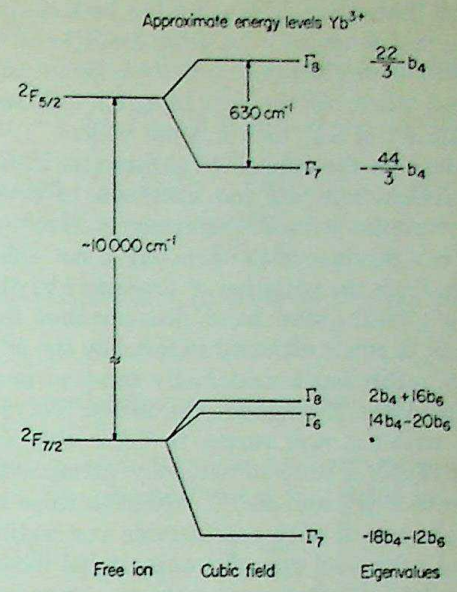


FIG. 5. Estimated cubic-field splitting of Yb^{3+} in CaF_2, assuming $A_4\langle r^4\rangle=-280$ cm^{-1} and $x=0.7$.

CONCLUSIONS

The cubic field in CaF_2 at trivalent rare-earth sites appears to be characterized by parameters in the order of magnitude

$$-A_4\langle r^4\rangle\sim250\text{–}430 \text{ cm}^{-1},$$
$$A_6\langle r^6\rangle\sim30\text{–}45 \text{ cm}^{-1}.$$

The extrapolation of data from divalent ions to the isoelectric trivalent ions is made subject to two assumptions based upon limited experimental evidence, namely that the ratio of the fourth- and sixth-order parameters of the trivalent ion is in the order of, or less than, that of the divalent ion, and that the trivalent ion parameters are larger than those of the divalent ion. The axial field parameters for nearest-neighbor charge compensation are more difficult to estimate from existing data. The $A_2^0\langle r^2\rangle$ term, however, probably has a magnitude of several hundred cm^{-1} for both tetragonal (n.n. F^- interstitial) and trigonal (O^{2-}) fields and hence produces a large perturbation on the cubic-field energy levels. The data are usually insufficient to evaluate the A_4^3, A_4^4, A_6^3, and A_6^4 terms for the noncubic crystal fields.

The parameters $A_n^m\langle r^n\rangle$ are expected to vary throughout the rare-earth series. If one assumes the A_n^m's are constants determined by appropriate lattice sums, then any variation of the crystal-field parameters would be due to $\langle r^n\rangle$. In the early work of Elliott and Stevens,[30] this variation was approximated by a multiplicative factor $(Z-55)^{-n/4}$, where Z is the atomic number. Application of this formula has not been universally

[66] P. P. Feofilov, Opt. i Spektroskopiya 5, 216 (1958).
[67] P. P. Pashinin and A. M. Prokhorov, *Paramagnetic Resonance* (Academic Press Inc., New York, 1963), p. 197.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

successful,[68] however. This point has been discussed by Freeman and Watson[69] who, from analyses of data for rare earths in trichlorides and ethylsulfates and using calculated values of the $\langle r^n \rangle$ integrals, conclude that the constancy of $A_n{}^m$ with Z is not valid.

The data on the cubic-field parameters in CaF_2 for most trivalent ions are too uncertain to derive any explicit variation with Z. Qualitatively, $A_4\langle r^4 \rangle$ appears to exhibit a decrease with increasing Z not unlike that predicted from the variation of Freeman and Watson's $\langle r^4 \rangle$. $A_6\langle r^6 \rangle$, on the other hand, does not show the large, monotonic decrease expected from solely the $\langle r^6 \rangle$ variation, but rather has a moderately constant or slightly increasing value throughout the series. The x values for the trivalent rare earths in cubic sites in CaF_2 given in Table I are estimated by extrapolating the behavior of $A_4\langle r^4 \rangle$ and $A_6\langle r^6 \rangle$. Although these x values must be accepted with considerable reservations, the predicted cubic-field ground states are not inconsistent with available paramagnetic resonance data, and they explain the absence of observed cubic-field resonances for Tb^{3+}, Ho^{3+}, and Tm^{3+}.

APPENDIX: CALCULATION OF CRYSTAL-FIELD MATRIX ELEMENTS USING 3-j AND 6-j SYMBOLS

The crystal-field potential V at a rare-earth ion site can be expanded as

$$V=\sum_{k,q} V_k{}^q=\sum_{k,q} A_k{}^q\langle r^k \rangle Y_k{}^q(\theta,\phi)\,, \qquad \text{(A1)}$$

where $Y_k{}^q(\theta,\phi)$ is the spherical harmonic of degree k and azimuthal quantum number q, $\langle r^k \rangle$ is the mean value of the kth power of the $4f$-electron radius, and the $A_k{}^q$'s are parameters related to the strength of the crystal field. Using the concept of operator equivalents, a general term in the crystal-field Hamiltonian may be rewritten as

$$\zeta_k A_k{}^q\langle r^k \rangle O_k{}^q\,, \qquad \text{(A2)}$$

where the ζ_k is Stevens' operator equivalent factor[32] α, β, γ for $k=2$, 4, 6, respectively, and the $O_k{}^q$'s are angular momentum operators, a complete list of which has been given by Orbach.[70] Tables of matrix elements of these operators within a J manifold and operator equivalent factors for the ground states of the trivalent rare-earth ions are located in various places in the literature. While this approach is convenient for the ground-state computations, when calculating the crystal-field matrix elements of excited states, crystal-field admixing between different J states, and in the treatment of spin-lattice relaxation where all values of k and q may be allowed, the more general approach given below which takes advantage of the tabulated 3-j and 6-j symbols of Rotenberg *et al.*[33] has been useful.

Elliott, Judd, and Runciman[71] have shown how the crystal-field matrix elements of a rare-earth ion in a configuration f^n can be calculated using the tensor operator techniques developed by Racah.[72] Their expression for the matrix elements of $U_q{}^k$, defined as $(4\pi/2k+1)^{1/2}Y_k{}^q$, rewritten in terms of 3-j symbols $(:::)$ and 6-j symbols $\{:::\}$ is

$$\begin{aligned}(f^n WUSLJJ_z|U_q{}^k|f^n W'U'S'L'J'J_z')&=\delta(SS')(-1)^{J_z+S+L'}[(2J+1)(2J'+1)]^{1/2}\\ &\times\begin{pmatrix} J & k & J' \\ -J_z & q & J_z' \end{pmatrix}\begin{Bmatrix} L & J & S \\ J' & L' & k \end{Bmatrix}\\ &\times(f^n WUSL\|U^k\|f^n W'U'S'L')\,, \qquad \text{(A3)}\end{aligned}$$

where for a rare-earth ion within a given configuration only $k=2$, 4, 6 need be considered. The reduced or double bar matrix element at the end of Eq. (A3) are

TABLE III. Reduced matrix elements of the lowest S-L term of the trivalent rare-earth ions. The first + or − sign corresponds to the first ion of each pair.

Ion	Ground term	$(\tau SL\|U^2\|\tau SL)$	$(\tau SL\|U^4\|\tau SL)$	$(\tau SL\|U^6\|\tau SL)$
Ce^{3+} Yb^{3+}	^2F	± 1	± 1	± 1
Pr^{3+} Tm^{3+}	^3H	$\pm\frac{1}{3}(11\times 13/2\times 7)^{1/2}$	$\mp\frac{2}{3}(13/7)^{1/2}$	$\pm\frac{1}{3}(5\times 17/7)^{1/2}$
Nd^{3+} Er^{3+}	^4I	$\pm(13/2\times 3\times 11)^{1/2}$	$\mp(2\times 13^2\times 17/3^2\times 11^2)^{1/2}$	$\pm(5/11)(17\times 19/3\times 7)^{1/2}$
Pm^{3+} Ho^{3+}	^5I	$\mp(13/2\times 3\times 11)^{1/2}$	$\pm(2\times 13^2\times 17/3^2\times 11^2)^{1/2}$	$\mp(5/11)(17\times 19/3\times 7)^{1/2}$
Sm^{3+} Dy^{3+}	^6H	$\mp\frac{1}{3}(11\times 13/2\times 7)^{1/2}$	$\pm\frac{2}{3}(13/7)^{1/2}$	$\mp\frac{1}{3}(5\times 17/7)^{1/2}$
Eu^{3+} Tb^{3+}	^7F	∓ 1	∓ 1	∓ 1

[68] M. J. D. Powell and R. Orbach, Proc. Phys. Soc. (London) **78**, 753 (1961).
[69] A. J. Freeman and R. E. Watson, Phys. Rev. **127**, 2058 (1962).
[70] R. Orbach, Proc. Roy. Soc. (London) **A264**, 458 (1961).
[71] J. P. Elliott, B. R. Judd, and W. A. Runciman, Proc. Roy. Soc. (London) **A240**, 509 (1957).
[72] G. Racah, Phys. Rev. **62**, 438 (1942); **63**, 367 (1943).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IV. Normalization factors $N_k{}^q$ for Eq. (A4). Numbers in column F are multiplicative factors common to all elements in the row. Other column headings denote q values.

$N_k{}^q$	F	0	1	2	3	4	5	6
$N_2{}^q$	$-4(7/3\times5)^{1/2}$	1	$1/(2\times3)^{1/2}$	$1/2(2\times3)^{1/2}$				
$N_4{}^q$	$8(2\times7/11)^{1/2}$	1	$1/4(5)^{1/2}$	$1/2(2\times5)^{1/2}$	$1/4(5\times7)^{1/2}$	$1/(2\times5\times7)^{1/2}$		
$N_6{}^q$	$-160(7/3\times11\times13)^{1/2}$	1	$1/2(2\times3\times7)^{1/2}$	$1/(3\times5\times7)^{1/2}$	$1/2(3\times5\times7)^{1/2}$	$1/3(2\times7)^{1/2}$	$1/6(7\times11)^{1/2}$	$1/(3\times7\times11)^{1/2}$

tabulated in Table III for the lowest L-S term of the trivalent rare-earth ions, where τ denotes the remaining quantum numbers required to specify the state. Many of the reduced matrix elements for the various ions are related by simple multiplicative factors derivable from the relationships discussed by Judd.[73] The reduced matrix elements of U^k for other L-S terms may be calculated using Racah's fractional parentage coefficients as shown by Elliott, Judd, and Runciman. Nielson and Koster[74] have recently tabulated these matrix elements for f^n configurations.

In Stevens' operator equivalence approach, certain normalization factors have been absorbed into the definition of the $A_k{}^q$. Thus if Eq. (A3) is used to evaluate crystal-field parameters and the $A_k{}^q$'s are to be compared with those derived elsewhere using operator equivalent factors, then (A2) should be replaced by

$$N_k{}^q A_k{}^q \langle r^k \rangle U_q{}^k. \tag{A4}$$

The normalization factors $N_k{}^q$ for the various k's and q's are listed in Table IV.

In terms of the 6-j symbols, the operator equivalent factors, ζ_k, for different J states are related by

$$\frac{(SLJ'\|\zeta_k\|SLJ')}{(SLJ\|\zeta_k\|SLJ)} = (-1)^{J'-J}\frac{2J'+1}{2J+1} \times\left[\frac{(2J'-k)!(2J+k+1)!}{(2J-k)!(2J'+k+1)!}\right]^{1/2} \times\begin{Bmatrix} J' & J' & k \\ L & L & S \end{Bmatrix}\Big/\begin{Bmatrix} J & J & k \\ L & L & S \end{Bmatrix}, \tag{A5}$$

where $k=2$, 4, 6 correspond to α, β, γ, respectively.

[73] B. R. Judd, Proc. Roy. Soc. (London) **A250**, 562 (1959).

[74] C. W. Nielson and George F. Koster, *Spectroscopic Coefficients for the p^n, d^n, and f^n Configurations*, (The M.I.T. Press, Massachusetts Institute of Technology, Cambridge, 1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Paramagnetic Resonance and Relaxation of Trivalent Rare-Earth Ions in Calcium Fluoride. II. Spin-Lattice Relaxation

R. W. Bierig, M. J. Weber, and S. I. Warshaw
Raytheon Research Division, Waltham, Massachusetts

(Received 3 January 1964)

The temperature dependence of the spin-lattice relaxation time T_1 of trivalent rare-earth impurities in CaF_2 has been measured at 9.6 kMc/sec using pulse saturation techniques. The rare earths investigated include Kramers ions (Ce^{3+}, Nd^{3+}, Dy^{3+}, Er^{3+}, Yb^{3+}), one non-Kramers ion (Tb^{3+}), and one S-state ion (Gd^{3+}). The relaxation of ions in lattice sites of cubic, tetragonal, and trigonal crystal-field symmetries have been studied and are found to be different. At liquid-helium temperatures the relaxation usually exhibits a T^{-1} temperature dependence; however, the magnitudes are approximately two orders-of-magnitude shorter than estimated single-phonon relaxation times based upon the orbit-lattice interaction. Measurements of the Ce^{3+} relaxation in five samples containing from 0.08 to 1.6% Ce combined with results from other ions show that T_1 decreases with increasing rare-earth concentration. Multiple exponential recoveries from saturation are observed at low temperatures and are discussed with respect to possible cross-relaxation within inhomogeneous spin systems and to other faster relaxing systems. At higher temperatures examples of both Raman and Orbach two-phonon relaxation processes are found. Calculations of two-phonon relaxation times using estimated values of the dynamic crystal fields, energy level structures, and ion wave functions yield coarse order-of-magnitude agreement with experimental values.

INTRODUCTION

SPIN-LATTICE relaxation of rare-earth ions has been the subject of considerable attention in the past few years, prompted by the discovery of the importance of two-step relaxation processes involving low-lying excited states[1] and the subsequent theoretical treatment by Orbach.[2] The relaxation of several different rare-earth ions has been measured in ethyl sulfate,[3] double nitrate,[3] and garnet crystals[4] and compared with theory with varying degrees of success. We have investigated the temperature dependence of the spin-lattice relaxation time, T_1, of trivalent rare-earth ions in CaF_2, including examples of Kramers (Ce^{3+}, Nd^{3+}, Dy^{3+}, Er^{3+}, Yb^{3+}), non-Kramers (Tb^{3+}) and S-state (Gd^{3+}) ions. This crystal is of additional interest since trivalent rare earths can enter the lattice substitutionally in sites having cubic, tetragonal, and trigonal crystal-field symmetry; thus one can compare the characteristics of the relaxation of ions in different symmetry sites in the same host lattice.

Calculations of the rate of spin-lattice relaxation and predictions of the relative importance of various relaxation mechanisms require knowledge of the energy level structures, ion wave functions, and strength of dynamic crystalline fields. Unfortunately, such information is not well established for most rare-earth ions in CaF_2. In the preceding paper[5] (henceforth referred to as I), available optical and electron paramagnetic resonance data were reviewed to obtain estimates of the cubic- and axial-crystal-field parameters and electronic energy levels. These results are used to estimate the magnitude and importance of various one- and two-phonon relaxation processes. Pertinent details of the resonance spectra, g values, etc., are also given in I.

The temperature dependence of T_1 for the above rare-earth ions has been measured from 2°K to as high a temperature as the time resolution (~ 3 μsec) and signal amplitude would permit. These results are presented following a brief review of the theory of spin-lattice relaxation in CaF_2 arising from modulation of the crystal field by thermal lattice vibrations and a discussion of the experimental approach. The T_1's at liquid-helium temperatures usually exhibit a T^{-1} temperature dependence characteristic of one-phonon relaxation processes; however, the magnitudes are about two orders-of-magnitude shorter than estimated relaxation times based upon the orbit-lattice interaction. In the case of Ce^{3+}, the low-temperature relaxation times are found to be a function of rare-earth concentration. Some possible mechanisms which may account for the enhanced and concentration-dependent relaxation rates are discussed qualitatively. At higher temperatures T_1 decreases rapidly with increasing temperature. In most cases the magnitude and temperature dependence of the relaxation times are in reasonable order-of-magnitude agreement with those for two-phonon Orbach and Raman processes calculated by using estimates of the crystal-field parameters and energy level splittings obtained in I.

THEORY

We are interested in the spin-lattice relaxation rate of the lowest resonant doublet of a rare-earth ion in a crystal field. Since no evidence of a phonon bottleneck[3] was observed in our experiments, we shall be concerned with the true spin-lattice rather than the lattice-bath relaxation rate. If the excited states are relatively far removed from the resonant ground doublet, then, following a short intense rf pulse applied to excite the spin

[1] C. B. P. Finn, R. Orbach, and W. P. Wolf, Proc. Phys. Soc. (London) 77, 261 (1961).
[2] R. Orbach, Proc. Roy. Soc. (London) A264, 458 (1961).
[3] P. L. Scott and C. D. Jeffries, Phys. Rev. 127, 32 (1962) and references therein.
[4] I. Svare and G. Seidel, *Paramagnetic Resonance* (Academic Press Inc., New York, 1963), p. 430.
[5] M. J. Weber and R. W. Bierig, Phys. Rev. 134, A1492 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

system from its equilibrium Boltzmann population, this two-level system will exhibit a simple exponential recovery with a characteristic time T_1.

There are several spin-lattice relaxation mechanisms which may be operative for rare-earth ions. The theory of rare-earth-ion relaxation has been developed by Orbach[2] and reviewed by Scott and Jeffries.[3] We shall only quote the final results here; the reader is referred to the above papers for more detail.

The simplest relaxation mechanism is a one-phonon direct process in which the spin-relaxation transition is accompanied by the absorption or emission of a single, energy-conserving phonon. The expression for T_1 due to the interaction between a spin and the fluctuating electric crystal field is given by

$$\frac{1}{T_{1D}}=\frac{3}{2\pi\rho v^5\hbar}\left(\frac{\delta}{\hbar}\right)^3|\langle a|\sum_{n,m}V_n{}^m|b\rangle|^2\coth(\delta/2kT), \quad (1)$$

where ρ is the mass density of the sample, v an average phonon velocity, δ the energy of the resonant transition between spin states $|a\rangle$ and $|b\rangle$, and the $V_n{}^m$'s are components of the dynamic crystal-field potential. The orbit-lattice interaction does not connect time conjugate states; consequently, if $|a\rangle$ and $|b\rangle$ are zero-order state functions, Eq. (1) is applicable only to non-Kramers ions, since for a Kramers ion the ground state is at least doubly degenerate and the crystal-field interaction between these states obviously vanishes. A Zeeman field may, however, admix states of the ground doublet with an excited state $|c\rangle$ at an energy Δ_c. Spin-lattice relaxation transitions are now allowed at a reduced rate governed by the magnitude of the Zeeman perturbation and given by

$$\frac{1}{T_{1D}}=\frac{3}{2\pi\rho v^5\hbar}\left(\frac{\delta}{\hbar}\right)^3\left(\frac{2\beta\Lambda}{\Delta^c}\right)^2\{|\langle a|\mathbf{H}\cdot\mathbf{J}|c\rangle\langle c|\sum_{n,m}V_n{}^m|b\rangle + \langle a|\sum_{n,m}V_n{}^m|c\rangle\langle c|\mathbf{H}\cdot\mathbf{J}|b\rangle|\}^2\coth(\delta/2kT), \quad (2)$$

where one must sum over all effective states $|c\rangle$. At temperatures above $\sim 1°$K, δ is usually $\ll 2kT$ and hence the direct process relaxation time, T_{1D}, is characterized by a T^{-1} temperature dependence.

An important relaxation mechanism for many rare-earth ions is the so-called Orbach process. This is essentially a two-step process involving real transitions to and from an excited state $|c\rangle$ accompanied by the creation and annihilation of two high-energy phonons. For this mode of resonance relaxation between the phonons and crystal-field splitting, one requires $\Delta_c<k\theta$ where θ is related to the cutoff frequency for phonons and is approximated by the Debye θ. The relaxation rate for this process is

$$\frac{1}{T_{1O}}=\frac{3}{2\pi\rho v^5\hbar}\left(\frac{\Delta_c}{\hbar}\right)^3\times\frac{|\sum_{\substack{n,n'\\m,m'}}\langle a|V_n{}^m|c\rangle\langle c|V_{n'}{}^{m'}|b\rangle|^2}{|\sum_{n,m}\langle a|V_n{}^m|c\rangle|^2+|\sum_{n',m'}\langle c|V_{n'}{}^{m'}|b\rangle|^2}\times\left(\exp\frac{\Delta_c}{kT}-1\right)^{-1} \quad (3)$$

and is characterized by an exponential temperature dependence at low temperatures, $T\ll(\Delta_c/k)$.

A two-phonon Raman relaxation process is one in which the energy of the spin transition is conserved during an inelastic phonon scattering. If the first-order term in the dynamic crystal-field interaction is used in second-order perturbation theory to calculate the transition probability, one finds that for non-Kramers ions at low temperatures, where $T\ll\theta$,

$$\frac{1}{T_{1R}}=\frac{9\times 6!}{4\rho^2\pi^3v^{10}\Delta_c{}^2}|\sum_{\substack{n,m\\n',m'}}\langle a|V_n{}^m|c\rangle\langle c|V_{n'}{}^{m'}|b\rangle|^2\times\left(\frac{kT}{\hbar}\right)^7. \quad (4)$$

For Kramers ions, again due to the consequences of time-reversal symmetry, there is a partial "Van Vleck cancellation"[6] such that in place of Eq. (4) one has

$$\frac{1}{T_{1R}}=\frac{9!\hbar^2}{\pi^3\rho^2v^{10}\Delta_c{}^2}|\sum_{\substack{n,m\\n',m'}}\langle a|V_n{}^m|c\rangle\langle c|V_{n'}{}^{m'}|b\rangle|^2\left(\frac{kT}{\hbar}\right)^9. \quad (5)$$

Another two-phonon Raman contribution arises from the use of the second-order term, in the variation of the crystal field caused by lattice vibrations, in first-order perturbation theory and is given by

$$\frac{1}{T_{1R}}=\frac{9\times 6!}{4\rho^2\pi^2v^{10}}|\sum_{n,m}\langle a|V_n{}^m|b\rangle|^2\left(\frac{kT}{\hbar}\right)^7, \quad (6)$$

for non-Kramers ions at low temperatures. This term is the same order of magnitude and additive to the result in Eq. (4). As for the direct process, the above interaction is active for Kramers ions only via the eigenstate admixing caused by the Zeeman field. This introduces a multiplicative factor, $(2\Lambda\beta H/\Delta_c)^2$, into Eq. (6). In general, the dominant Raman relaxation process for Kramers ions is given by Eq. (5) with a T_{1R} proportional to T^{-9} compared to the T^{-7} temperature dependence for non-Kramers ions.

When a four-level Γ_8 quartet is lowest or in the case of an S-state ion, the effective spin system may not.

[6] J. H. Van Vleck, Phys. Rev. **57**, 426 (1940).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

consist simply of a single isolated doublet. The relaxation modes in such multilevel systems are, in general, more complex. At low temperatures the dominant relaxation process between pairs of levels may occur via multistep direct processes and exhibit a transient time dependence proportional to a linear combination of several exponential terms.[7] Also, at the low-temperature end of the Raman relaxation region, a T_1 mechanism proportional to T^{-5} may be dominant.[8] These cases, however, are somewhat specialized examples of rare-earth-ion relaxation.

To evaluate T_1 from the expressions in Eqs. (1)–(6) the matrix elements of V_n^m are required. For the rare earths, one need only consider n even and ≤ 6. Which components of V_n^m enter into the calculation of the relaxation time in a given crystalline environment can be determined from a normal-mode analysis[2] of the vibrations of the paramagnetic ion and the nearest-neighbor surrounding charges, these being assumed to generate the principal orbit-lattice coupling. The vibrational modes transform as some combination of the irreducible representations of the crystal symmetry group. For CaF_2, which has cubic symmetry, the vibrational motion of the rare-earth ion and the eight nearest-neighbor F^- ions transform according to the following irreducible representations of the O_h group:

$$\Gamma_{\rm vib}(O_h)=A_{1g}+A_{2u}+E_g+E_u+2T_{1u}+2T_{2g}+T_{2u}, \quad (7)$$

where for transitions within a given electronic configuration only the even vibrations need be considered and the translational and rotational terms have been subtracted out. Similarly, for the tetragonal and trigonal site symmetries described previously,[5]

$$\Gamma_{\rm vib}(C_{4v})=5A_1+A_2+2B_1+4B_2+6E, \quad (8)$$

and

$$\Gamma_{\rm vib}(C_{3v})=6A_1+A_2+7E. \quad (9)$$

Comparison of the irreducible representations in (7), (8), and (9) and those according to which the V_n^m's transform, shows that in all cases the vibrational motion is factored into a number of representations such that all V_n^m's are allowed. Thus there is no reduction and simplification in the number of terms which must be considered in applying the above theoretical expressions to the evaluation of T_1 for ions in the different symmetry sites in CaF_2.

Attempts to make explicit calculations of spin-lattice relaxation times are plagued by the necessity of making numerous assumptions and approximations. Given expressions for T_1 as in Eqs. (1)–(6), which are derived using a simplified treatment of the lattice vibrations, estimates must be made of the dynamic crystal-field components. Two approximations which have been used[2,3] in making such estimates are that (1) the orbit-lattice interaction parameters may be approximated by static crystal-field parameters, i.e., $|(x\partial A_n^m/\partial x)| \approx |A_n^m(x_0)|$, where x is a ligand coordinate displacement, and (2) $|A_n^m| \approx |A_n^0|$. Although the latter rule was based upon empirical data, it is not universally satisfied. Other approaches[9] include estimates of the strength of the orbit-lattice parameters obtained from uniaxial stress and ultrasonic spin-phonon absorption experiments. Convincing calculations of the parameters would require knowledge of the nature of the ligand fields, shielding effects, and vibrational motion. In the following we shall consider those combinations of matrix elements of the orbit-lattice interaction which appear in the expressions for the direct, Orbach, and Raman T_1's as parameters to be determined experimentally. These are compared later with rough order-of-magnitude estimates using the above approximations and the crystal-field parameters and energy level splittings found in I.

For a resonance frequency of 9.6 kMc/sec and taking $\rho=3.18$ g/cm^3 and a mean phonon velocity, $v=4\times10^5$ cm/sec for CaF_2, T_1 for the direct process in Eq. (1) becomes

$$T_{1D}=(1.5/TM_D^2)(\sec), \quad (10)$$

where for non-Kramers ions,

$$M_D^2\equiv|\langle a|\sum_{n,m}V_n^m|b\rangle|^2(\text{cm}^{-2}), \quad (11)$$

and for Kramers ions,

$$M_D^2\equiv(2\beta\Lambda/\Delta_c)^2\{|\langle a|\mathbf{H}\cdot\mathbf{J}|c\rangle\langle c|\sum_{n,m}V_n^m|b\rangle + \langle a|\sum_{n,m}V_n^m|c\rangle\langle c|\mathbf{H}\cdot\mathbf{J}|b\rangle|\}^2(\text{cm}^{-2}). \quad (12)$$

For the Orbach process, from Eq. (3),

$$T_{1O}\approx\frac{\exp(\Delta_c/kT)}{4\Delta_c^3(\text{cm}^{-3})M_O^2}(\sec), \quad (13)$$

where

$$M_O^2\equiv\frac{|\sum_{\substack{n,m\\n',m'}}\langle a|V_n^m|c\rangle\langle c|V_{n'}^{m'}|b\rangle|^2}{|\sum_{n,m}\langle a|V_n^m|c\rangle|^2+|\sum_{n'm'}\langle c|V_{n'}^{m'}|b\rangle|^2}(\text{cm}^{-2}). \quad (14)$$

A summation is to be made over all effective excited states $|c\rangle$ in (12) and (14). For Kramers salts, the T^{-9} Raman relaxation mechanism in Eq. (5) is usually dominant and reduces to

$$T_{1R}\approx(10^7\Delta^4/T^9M_R^4)(\sec), \quad (15)$$

where

$$M_R^2\equiv|\sum_{\substack{n,m\\n',m'}}\langle a|V_n^m|c\rangle\langle c|V_{n'}^{m'}|b\rangle|(\text{cm}^{-2}). \quad (16)$$

[7] R. W. Bierig and M. J. Weber, Phys. Rev. 132, 164 (1963).
[8] R. Orbach and M. Blume, Phys. Rev. Letters 8, 478 (1962).
[9] M. Blume and R. Orbach, Phys. Rev. 127, 1587 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Scott and Jeffries[3] found from the actual computation of sums of the type $\sum_{n,m}|\langle a|V_n{}^m|c\rangle|^2$ (assuming the individual terms to be incoherent), that for several rare earths these sums were of the same order of magnitude. Thus one might expect that $M_D{}^2$ (allowing for Zeeman admixing), $M_O{}^2$, and $M_R{}^2$ would not differ greatly. Exceptions are always possible, however, and would depend upon the actual state wave functions. Note, also, that if the magnitudes of $|\sum\langle a|V_n{}^m|c\rangle|^2$ and $|\sum\langle c|V_n{}^m|b\rangle|^2$ in Eq. (14) are very different, $M_O{}^2$ will be determined by the smaller of the two.

EXPERIMENTAL

A schematic representation of the pulse saturation X-band spectrometer used in measuring spin-lattice relaxation is shown in Fig. 1. The principles of operation are straightforward. A VA 153 reflex klystron is used as a main rf power source. The klystron environment is a temperature-stabilized oil bath, an important improvement for obtaining good low-noise klystron operation. The VA 153 is frequency stabilized to the sample cavity resonance. Relatively high-power rf pulses are generated by two states of X-band amplification obtained using VA 824B klystron amplifiers driven by the VA 153. The rf pulse width can be varied from 2–500 μsec and peak output power is $\approx$20 W. Klystron amplifiers have the advantage that their output pulse is monochromatic to the extent that the driving element is frequency stabilized, hence there are no high-peak power FM signals reflected from the sample cavity to interfere with detection of the low-level transient relaxation signals. A portion of the VA 153 output is used as a low-level cw monitor of the sample cavity reflection coefficient. The cw power level incident on the sample cavity is normally kept between 0.1 and 10 μW.

The transient signals generated by relaxation of the spin system, following a high-power saturating pulse, are detected in a standard mixer; subsequent amplification is performed at 30 mc, the difference frequency at which the local oscillator is stabilized. The transient relaxation signals are displayed on a Tektronix oscilloscope and photographed. The VA 153 stabilization loop has a variable bandwidth (1–100 kc/sec) to permit rapid correction of the output frequency so that the sample cavity may be transiently tracked. If this is not done and care is not exercised so that the sampling frequency is at the center of cavity resonance, a mixture of time-dependent absorption and dispersion may be detected. The decay times of the real and imaginary components of the reflection coefficient are not necessarily the same, hence considerable confusion in the data may occur if this admixture is not taken into account. An alternative to rapid stabilization is to measure the decay time of the observed transient signal far out in the tail where the dispersion is small.

Diminution of pulse power incident on the mixer is accomplished by synchronously switching a double-diode switch to reflect pulse power back into an absorbing isolator. The time resolution of the spectrometer, which is a few microseconds, is determined by the "turn-off" time of the klystron pulse amplifiers and by the recovery time of the preamplifier "front end" from saturation.

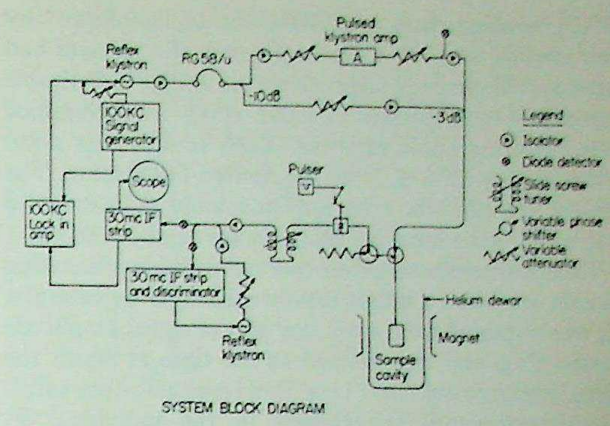


FIG. 1. Schematic diagram of pulse saturation X-band spectrometer.

The sample cavity resonates in the TE011 circular electric mode and has a loaded $Q\approx$8000–10 000. The sample under study is mounted axially in the resonant chamber on a sapphire tube. At low temperatures the sapphire provides excellent thermal coupling to the cavity which, in turn, is in thermal contact with the helium bath via helium exchange gas. The cavity mass acts as the thermal reservoir to which the spin system relaxes. The cavity temperature is carefully monitored by a set of 3 thermometers; two 47-Ω carbon resistors and a platinum resistance thermometer. The carbon resistors were individually calibrated between 1.5 and 20°K; the platinum resistance thermometer was calibrated from 4.2–300°K. These three thermometers were used together with a gold-cobalt differential thermocouple connected between one of the carbon resistors and the sample mount to verify that the sample temperature did indeed track the cavity temperature in the range 2–50°K. The maximum temperature differential was found to be less than 0.1°K in this temperature range. This was important, since all relaxation measurements at temperatures above 4.2°K were made by simply allowing the cavity structure to warm toward 77°K. The total time available for measurements between 4.2 and 50°K is typically 3.5 h.

Most of the CaF_2 crystals used in these experiments were grown in this laboratory by the Czochralski method[10a] or by a modified Stockbarger[10b] technique. The melts were contained in molybdenum crucibles and the growth ambient consisted of dry argon gas [dew point <-100°F]. Other samples were purchased from Optovac, Inc. The crystals were annealed by sealing them in evacuated quartz ampoules [10^{-6} mm Hg at

[10] (a) K. Nassau, J. Appl. Phys. 32, 1820 (1961). (b) D. C. Stockbarger, J. Opt. Soc. Am. 39, 731 (1949).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

20°C], heating them to 1200°C for 12 h, followed by slow cooling at ~20°C/h. In general, the samples had rare-earth dopings of the order of 0.1 wt.%; undesirable impurities were present $\lesssim$0.001 wt.% as determined from spectrographic analysis. Each sample was x-ray oriented to within ±2° of the desired orientation. The typical sample size was approximately 0.1×0.1×0.3 cm, which gave a cavity filling factor of about 0.001.

Except where noted, the observed transient relaxation signals exhibited a single exponential time dependence. T_1 was obtained from a semilog plot of signal amplitude versus time and was defined as the time at which the signal had recovered to $(1-e^{-1})$ of its equilibrium value. The experimental uncertainty in the individual T_1 determinations is typically ~10%. The electron paramagnetic resonance (EPR) spectrum was examined as a function of cw power level to insure that the sampling signal intensity was always at least 20 dB below the power level at which steady-state saturation could be observed. For a few of the samples (Ce^{3+} and Gd^{3+}) the incident power had to be reduced to a few tenths of a microwatt. The saturating pulse width τ_p was always adjusted such that $\tau_p \lesssim T_1$.

RESULTS

Cerium

Spin-lattice relaxation of Ce^{3+} in tetragonal sites in CaF_2 was investigated in the temperature range 2–15°K for five samples having different Ce concentrations ranging from 0.08 to 1.6 wt.% cerium. In all samples the tetragonal site resonances were the most intense. Additional lines were usually present in the resonance spectrum, particularly at the higher Ce concentrations (see Fig. 2 of I); however, they were generally too weak for relaxation measurements. T_1 measurements were made with the magnetic field H applied in a (100) crystallographic plane.

Figure 2 shows the over-all temperature dependence of T_1 for several different Ce^{3+} concentrations. The transient recovery from saturation of the signal from the 1.6% sample displayed a single exponential time dependence with a characteristic time which was approximately temperature-independent at low temperatures. The recovery of the other samples, however, could not be fitted by a single exponential term at low temperatures but rather by a sum of two or three exponential terms. An initial rapid recovery was usually observed which was more or less temperature-independent but concentration-dependent. This component had a characteristic time varying from ≈100 μsec for the 0.34% sample to ≈300 μsec for the 0.08 and 0.18% samples. When saturation pulses of several hundred microseconds duration were applied, this component was diminished. Thus it appears to be associated with the establishment of internal equilibrium within the spin system[11] and not

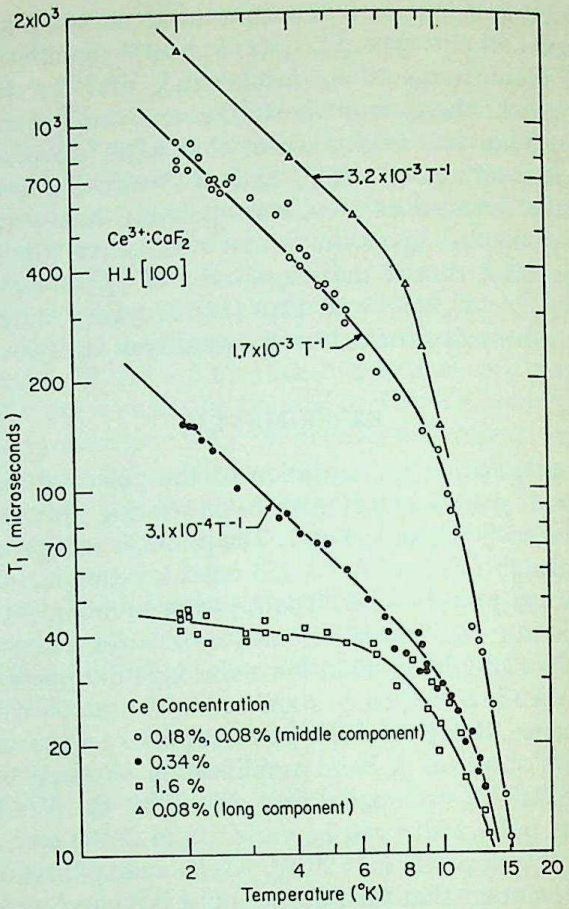


FIG. 2. Temperature dependence of relaxation time T_1 of Ce^{3+} in tetragonal sites in CaF_2 for several different Ce concentrations.

spin-lattice relaxation. Following it was a temperature-dependent recovery composed of one or two exponential terms. The dominant temperature-dependent component of the recovery is concentration-dependent and is plotted as T_1 in Fig. 2. A final rather weak temperature-dependent component of the recovery is also plotted in Fig. 2 for the 0.08% sample. Except for the most concentrated 1.6% sample, the relaxation in the low-temperature region is characterized by a T^{-1} temperature dependence and a concentration-dependent magnitude.

At higher temperatures, $T>10$°K, the data for the different samples tend to converge to a value independent of the cerium concentration. The asymptotic T_1 temperature dependence can be fitted to a T^{-9} law for Raman relaxation of a Kramers ion using $T_{1R}\approx 4\times 10^5\ T^{-9}$ sec. The data can also be fitted to an exponential law with $\Delta\sim 150$ cm^{-1}; however, the fit is poorer.

For several of the rare-earth ions studied it is difficult to establish an unambiguous temperature dependence at high temperatures since the magnitude of T_1, after subtracting out the T^{-1} dependence at low temperatures is not known accurately over a sufficiently broad range of temperature.

If the asymptotic T_1 temperature dependence is fitted to be a T^{-9} law for Raman relaxation, we find from Eq. (15) that $M_R^2\approx 5\Delta^2$. The first excited state for Ce^{3+}

[11] K. D. Bowers and W. B. Mimms, Phys. Rev. 115, 285 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in tetragonal sites should be of form $|\pm\frac{1}{2}\rangle$. Sums of the type $\sum|\langle a|V_n^m|c\rangle|^2$ connecting this state to the ground state have values $\approx 1.5\times10^4$ cm^{-2} for $A_2^m\langle r^2\rangle\sim200$ cm^{-1} and $A_4^m\langle r^4\rangle\sim400$ cm^{-1}. (More correct wave functions, including admixtures of states from $^2F_{7/2}$, would introduce terms involving V_6^m; however, they are not expected to change the order of magnitude of the above result.) An M_R^2 of this magnitude, however, implies a small Δ which is inconsistent with the dominance of Raman relaxation. If we consider Orbach relaxation via an excited state at ~150 cm^{-1}, we find from the data and Eq. (13), $M_O^2\sim10^4$ cm^{-2}. Although the fit to an exponential temperature dependence is poorer, the magnitude agreement is better. A comparable Raman relaxation rate for $\Delta\sim150$ cm^{-1} requires orbit-lattice matrix elements $M_R^2\sim10^5$ cm^{-1}. Although there is a large uncertainty in our estimates of the V_n^m's, the above values of M_O^2 and M_R^2 are not unreasonable. Also an excited-state separation of a few hundred cm^{-1} for Ce^{3+} is possible if the tetragonal fields are as large as they appear to be for other rare earths in CaF_2.[5] Thus the above estimates are inconclusive in establishing whether the high-temperature relaxation is Raman or Orbach. The former, however, requires the larger orbit-lattice interaction.

The angular variation of T_1 with magnetic field direction in a (100) plane is shown in Fig. 3. The measurements were made using a 0.34% Ce sample at 2°K. The same ratio of T_1's for the $g_{\parallel}$ and $g_{\perp}$ resonances, i.e., for the magnetic field H applied along [100] and [010], was observed from the 0.18% sample, while for the 1.6% sample the two T_1's were equal. Since the $g_{\parallel}$ resonance is weaker than the $g_{\perp}$ resonance, the relaxation could not be measured in the former orientation for the 0.08% sample.

If the first excited tetragonal state is $|\pm\frac{1}{2}\rangle$, direct relaxation should be highly anisotropic since there is no Zeeman admixing for $\theta=0$. The over-all shape of the curve in Fig. 3, however, does not correspond exactly to that expected for an excited state of either $(\pm\frac{1}{2})$ or $(\pm\frac{5}{2}, \mp\frac{3}{2})$ and may be related to the concentration-dependent effects. The dip in T_1 near $\theta=80°$ occurs in the vicinity of the crossover of the resonances from Ce^{3+} in sites of tetragonal and trigonal symmetries. This suggests the possibility of cross-relaxation effects, although such effects are not observed when cubic and tetragonal Yb^{3+} resonances are overlapped, as we report later. Within experimental error, the relaxation at this orientation exhibited a simple exponential recovery.

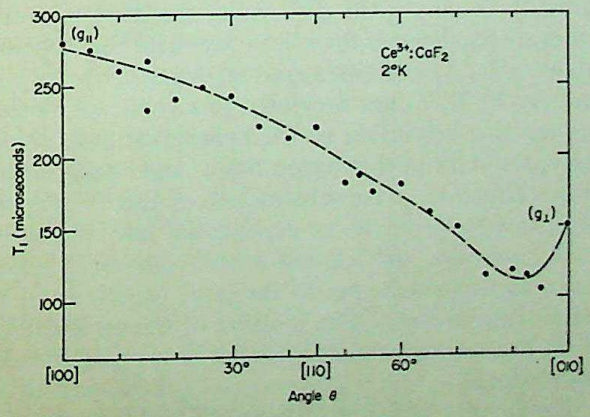


FIG. 3. Angular variation of T_1 for Ce^{3+} in tetragonal sites in CaF_2. The magnetic field is applied at an angle θ with respect to a [100] axis in a (100) plane.

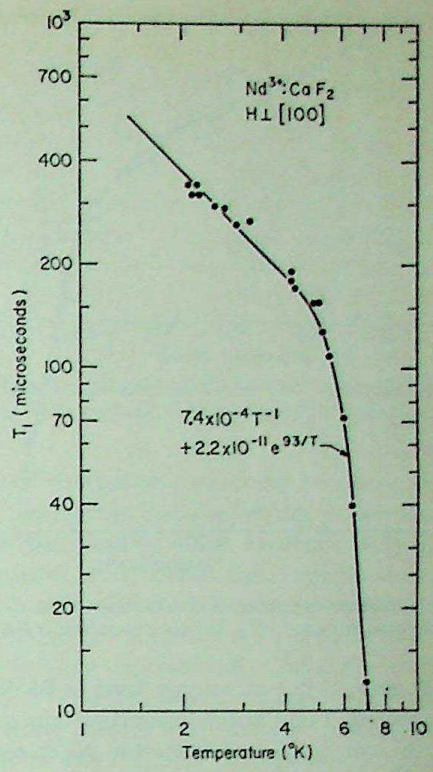


FIG. 4. Temperature dependence of relaxation time T_1 of Nd^{3+} in tetragonal sites in CaF_2. Nd concentration is 0.28%.

Neodymium

The temperature dependence of T_1 for Nd^{3+} in tetragonal sites is shown in Fig. 4, where the $g_{\perp}$ resonance for $H\perp[100]$ has been measured. Several weak broad ($\sim$40–100 G) lines whose origin and nature were not established were present in the resonance spectrum. At 2°K the relaxation time of the $g_{\parallel}$ resonance is about twice as long as that of the $g_{\perp}$ resonance. At low temperatures T_1 exhibits a T^{-1} dependence. At the higher temperatures, $T>4$°K, a relaxation mechanism decreasing more rapidly than T^{-9} with increasing temperature begins to dominate. It can be fitted to an exponential temperature dependence given by $T_1=2.2\times10^{-11}\exp(93/T)$ sec. This suggests relaxation via an excited state or states effectively at ≈65 cm^{-1} above the ground state. The optical fluorescence data in I indicates a first excited state at ≈90 cm^{-1}. Within experimental error the two results are probably compatible. A tetragonal crystal-field component $A_2\langle r^2\rangle$ of a few hundred cm^{-1} would greatly perturb the two cubic-field Γ_8 quartets at

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

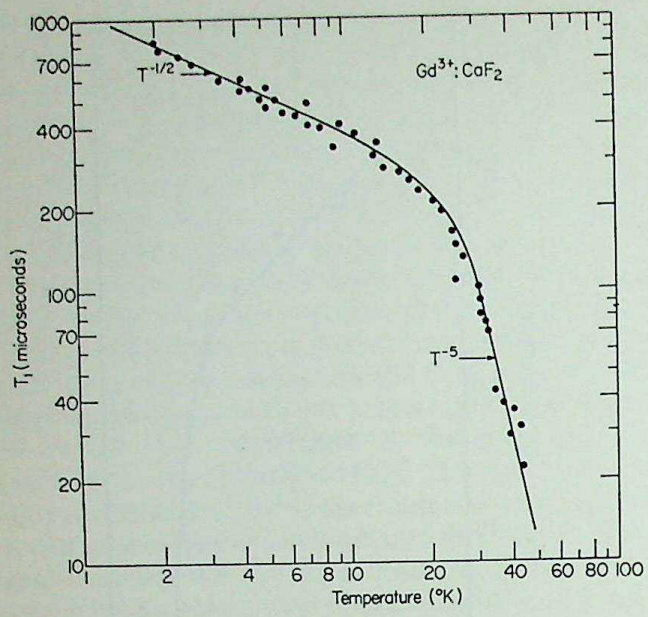


FIG. 5. Temperature dependence of relaxation time T_1 of Gd^{3+} in tetragonal sites in CaF_2. Gd concentration is 0.03%.

0 and 180 cm^{-1}, hence an energy level at 65–90 cm^{-1} is not unexpected. For Orbach relaxation via an excited state $\Delta=65$ cm^{-1}, we find from the Nd^{3+} results and Eq. (13), $M_0^2\approx 4\times 10^4$ cm^{-2}.

Gadolinium

The relaxation of Gd^{3+} ions in tetragonal sites was investigated. At liquid-helium temperatures the relaxation times of all transitions observable at X-band frequencies were essentially equal. Extensive measurements from 2 to 45°K were made of the temperature dependence of T_1 for the $\frac{1}{2}\leftrightarrow-\frac{1}{2}$ transition with $H\|[100]$ and are shown in Fig. 5. The over-all temperature dependence of the T_1 data can be fitted by $T_1=1.1\times10^{-3}$ $T^{-1/2}+4\times10^3$ T^{-5} sec. The transient recoveries were essentially single exponentials. Gadolinium ions were also present in sites of trigonal symmetry in the sample used; however, the relaxation was not examined.

Since an S state is unperturbed by a crystal field in first order, the crystal-field splittings and relaxation rates are usually very small. This is indeed observed for Gd:CaF_2, although the relaxation times below $\sim$10°K are unexpectedly short, again, as in the case of other rare-earth ions in CaF_2, less than 1 msec. This rapid T_1 and its $T^{-1/2}$ temperature dependence suggests that the low-temperature relaxation rate is not associated with the normal direct process and orbit-lattice interactions. The relaxation times at higher temperatures are longer than those for other non-S-state ions. The T^{-5} dependence may arise from two-phonon relaxation processes involving excited states within the $S_{7/2}$ multiplet[8] or relaxation of time conjugate states associated with a lattice defect site.[12] Since the measurements were made in a range where $T\lesssim 0.1\theta$ and θ_D for CaF_2 is $\approx$480°K, a $T_1\propto T^{-5}$ region is not expected at these temperatures for normal Raman relaxation processes.

Orbach and Blume[8] have shown that for a multilevel spin system at low temperatures, Raman relaxation via excited states within the multiplet will be proportional to T^{-5} and can be more important than T^{-7} and T^{-9} processes involving higher excited states. In general, when the excited-state energy is very small compared to the energy of the most important phonons for relaxation, the denominator of the second-order Raman process is $\approx(\hbar\omega_p)^2$ rather than Δ^2 as in Eq. (4). Thus one has in place of Eq. (4),

$$\frac{1}{T_{1R}}=\frac{9}{4\pi^3\rho^2v^{10}}\frac{4!}{\hbar^2}\left(\frac{kT}{\hbar}\right)^5 M_R{}^4. \tag{17}$$

$M_R{}^4$ is again given by Eq. (16), but now $|c\rangle$ is a state within the ground multiplet.

The energy levels for Gd^{3+} in a tetragonal field in CaF_2 and for a resonance frequency of 9.6 kMc/sec for the $\frac{1}{2}\leftrightarrow-\frac{1}{2}$ transition can be derived from the parameters given by Vinokurov *et al.* (see I). They form a multilevel system in which the Orbach-Blume mechanism may be operative. If the effective dynamic crystal-field interactions have the same relative importance as the static crystal-field parameters in the spin Hamiltonian, then the $b_2{}^m$ terms should dominate and Raman relaxation would proceed principally via the $|\pm\frac{3}{2}\rangle$ states. Comparison of Eq. (17) with the experimental T_1's for Gd:CaF_2 yields $M_R{}^2\approx 1.4\times10^3$ cm^{-2}. This value while smaller than for Kramers ions, is, however, several orders of magnitude larger than that predicted using the static crystal-field parameters.

Terbium

Trivalent terbium is the only non-Kramers ion studied. A plot of T_1 versus T for an $H\perp[111]$ sample orientation is shown in Fig. 6. The relaxation times of each of the four Tb^{159} hyperfine components were essentially equal. In the T^{-1} region the Tb^{3+} relaxation times are much faster than those found for the other ions. Above 4°K, T_1 exhibits a gradual departure from a T^{-1} law. By 10°K, T_1 has decreased to a value approaching the limiting resolution of the apparatus; and, accordingly, the data in this region has a larger experimental error. The data, after substraction of the T^{-1} dependence, can be fitted to an exponential law with $\Delta\sim20$ cm^{-1}; however, one cannot attach too much significance to this result due to the large experimental uncertainties involved. The location of the excited levels of $^7F_6$ in a tetragonal field have not been reported to our knowledge.

The angular variation of T_1 at 2°K and for a magnetic field applied in a (111) plane is shown in Fig. 7. Broad anisotropic resonances previously reported by Forrester

[12] J. G. Castle, D. W. Feldman, P. G. Klemens, and R. A. Weeks, Phys. Rev. **130**, 577 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and Hempstead[13] in Tb:CaF_2 were also observed in the sample used. The presence of Tb^{3+} in sites of trigonal symmetry, which are not observable at X-band frequencies, was not established.

Terbium, being a non-Kramers ion, is expected to relax faster at low temperatures than Kramers ions, which is indeed observed. The M_D^2 determined from Eq. (10) is $\approx 3\times10^4$ cm^{-2}. The resonant doublet, however, is composed principally of $J_z=\pm6$ states which are not connected by the orbit-lattice interaction. Relaxation may be allowed, for example, by admixing of $J_z=\pm2$ states. This small admixture would reduce the value of M_D^2 relative to that for other ions.

The above measurements were made in a sample containing 0.5% Tb. The relaxation from a 1.4% Tb sample was also measured between 2–4°K, where T_1 ~5 μsec, independent of temperature.

Dysprosium

Two different samples of dysprosium-doped CaF_2 ($\sim$0.1–0.2% Dy) each containing Dy^{3+} in sites of cubic tetragonal, and trigonal symmetries were examined. No measurable relaxation recoveries were observed from any of the resonances for temperatures 2–4°K. Based upon the limiting resolution of the apparatus, this result suggests that the relaxation times of the Dy^{3+} resonances in the samples studied are $\gtrsim3$ μsec.

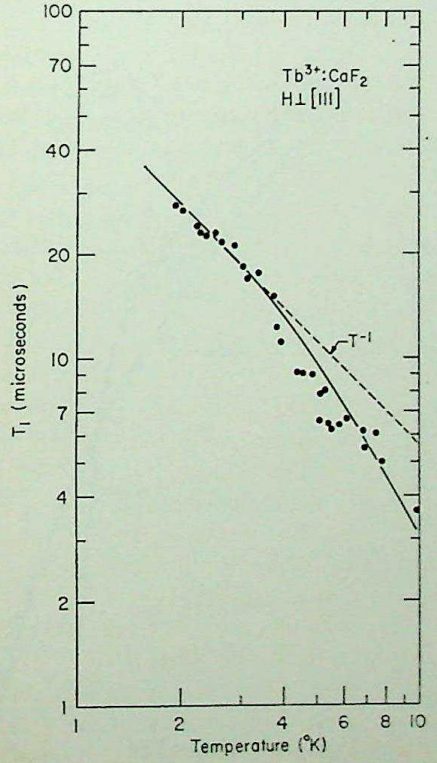


FIG. 6. Temperature dependence of relaxation time T_1 of Tb^{3+} in tetragonal sites in CaF_2. Tb concentration is 0.5%.

[13] P. A. Forrester and C. F. Hempstead, Phys. Rev. **126**, 923 (1962).

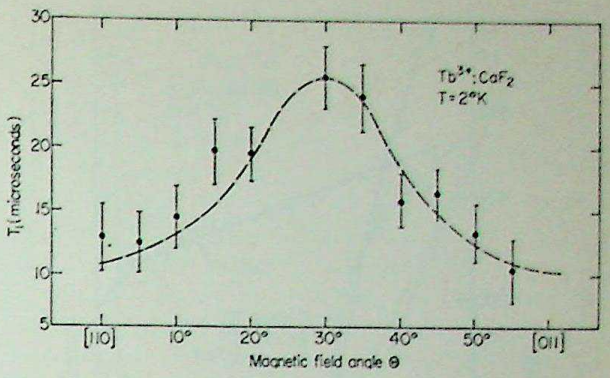


FIG. 7. Angular variation of T_1 for Tb^{3+} in tetragonal sites in CaF_2. The magnetic field is applied in a (111) plane.

The Dy^{3+} relaxation rates may be faster than those of other rare earths because of the presence of a relatively low-lying excited state. In a cubic field the relaxation of pairs of levels within the ground-state Γ_8 quartet may occur via an Orbach process involving the Γ_7 excited state at $\Delta\approx8.5$ cm^{-1}. A T_1 of 1 μsec, which would account for the unresolved relaxation transient, is predicted at 4°K for such a process if $M_O^2\approx10^4$ cm^{-2}. Corresponding times for Raman relaxation require $M_R^2\approx4\times10^5$ cm^{-2}. If this process was dominant, however, lifetime broadening of the Γ_8 resonances should be evident by $T\approx10$°K, which is not observed. At $T=2$°K, the major relaxation mechanism may be two-step relaxation processes involving states within the Γ_8 quartet, as discussed elsewhere.[7]

The relaxation rate of the cubic field Γ_7 excited state would also be fast if the strength of the orbit-lattice interaction is as strong as required above, the temperature dependence of an Orbach process changing from a constant at low temperatures, $T\ll\Delta$, to a T^{-1} law at high temperatures. At temperatures $\lesssim15$°K, lifetime broadening of the Γ_7 resonance linewidth is observed. From the estimated magnitude and slow temperature variation of the Γ_7 Orbach process, the broadening probably arises from the dominance of Raman relaxation at these temperatures. From Eq. (15), a T_{1R} of 10^{-3} sec, and hence line broadening, is predicted at $T=15$°K if $M_R^2\approx4\times10^4$ cm^{-2}.

Erbium

The temperature dependence of the relaxation of Er^{3+} in sites of tetragonal (I) and trigonal (II) symmetry are recorded in Fig. 8. The position and relative intensities of the resonances in the low-field portion of the spectrum of the 1.0% Er^{3+} sample used are shown in Fig. 5 of I for $H\|[110]$. Due to the strong overlap and relative intensities of the tetragonal (I) and cubic resonances, only a common transient decay was observed when the magnetic field was set in the region $\approx$950–1070 Oe, the amplitude varying as the magnetic field was adjusted to line center. The transient signal, which appears to be dominated by the tetragonal (I)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

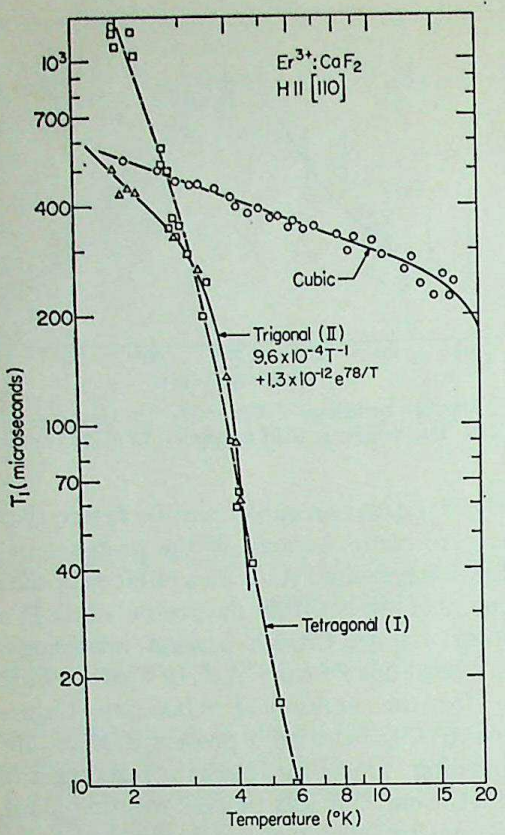


FIG. 8. Temperature dependence of relaxation time T_1 of Er^{3+} in tetragonal and trigonal sites in CaF_2. Er concentration is 1.0%. Cubic T_1 is from another sample (see text).

resonance, is characterized by a temperature-independent component ($\sim$100 μsec) and a strongly temperature-dependent component which is plotted as T_1 (tetragonal) in Fig. 8. The variation of T_1 with angle of the magnetic field in a (100) plane is at most 20%; there is no measurable variation of the initial decay component. At low temperatures, $\sim$2°K, a third decay component is evident having a characteristic time $\sim$2–5 msec. The amplitude of this signal, however, is so small that the experimental uncertainty in the decay time is very large.

The high-temperature portion of the tetragonal (I) T_1 versus T curve can be fitted by an exponential law given by $T_1\approx10^{-7}\exp(27/T)$ sec. At low temperatures the data deviates from the exponential curve; however, the data is not sufficient to establish a reliable temperature dependence for the competing relaxation process, which might be expected to have a $1/T$ law. An effective tetragonal excited state at $\sim$27°K from relaxation data is not unreasonable in view of the crystal-field splittings in I and the excited state at $\sim$35°K reported by Baker *et al.* from measurements of the temperature dependence of the intensity of the tetragonal excited-state resonance. Assuming an Orbach process is operative, the Mo^2 derived from the above result and Eq. (13) is $\approx10^3$ cm^{-2}. Line broadening of the cubic-tetragonal (I) group of resonances occurs for $T\gtrsim30$°K; by 40°K the individual lines are no longer resolvable.

The trigonal (II) relaxation was measured with the magnetic field applied in a [110] direction to separate the g_1 trigonal resonance as far as possible from the other spectral lines. The results in Fig. 8 show a T^{-1} region and a rapidly decreasing T_1 at higher temperatures which can be fitted to an exponential law. The over-all temperature dependence can be fitted by

$$T_1=9.6\times10^{-4}T^{-1}+1.3\times10^{-12}\exp(78/T)\text{ sec.}$$

Comparing this result with an Orbach process in Eq. (13) yields $Mo^2\approx10^6$ cm^{-2}. The linewidth of the trigonal (II) resonance increased rapidly for temperatures $\lesssim$25°K and became unobservable above 30°K indicative of lifetime broadening and $T_1\lesssim10^{-9}$ sec.

T_1 measurements were made earlier[14] in another Er:CaF_2 crystal which was less heavily doped than the above sample and in which Er^{3+} ions in cubic sites were approximately twice as numerous as those in tetragonal (I) sites. T_1 data for ions in trigonal (II) sites, which were also present, agreed with the data presented above. Data taken in the vicinity of the isotropic cubic resonance is shown in Fig. 8. For this sample the tetragonal (I) resonances could not be sufficiently well isolated to yield unambiguous results. Between 2 and 15°K, T_1 for the cubic resonance exhibited very little temperature dependence and can be fitted by $T_1\propto T^{-n}$, where $n\approx\frac{1}{3}$. This rather unexpected result was reproduced under varied experimental conditions and appears to be

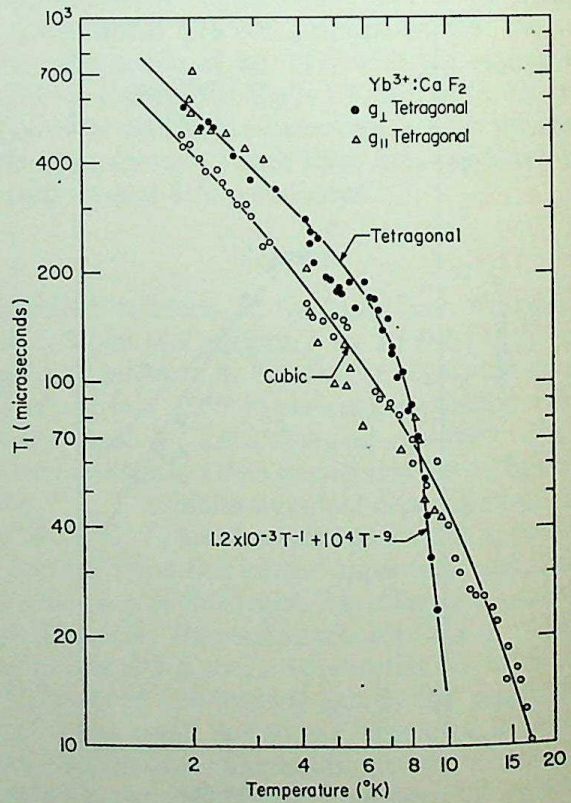


FIG. 9. Temperature dependence of relaxation time T_1 of Yb^{3+} in cubic and tetragonal sites in CaF_2. Yb concentration is 0.17%.

[14] M. J. Weber and R. W. Bierig, Bull. Am. Phys. Soc. 8, 259 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

intrinsic to the sample. Since there should be cubic-field excited states within 100 cm^{-1} of the ground state, such long T_1's for $T>10°K$ are very unusual.

The relaxation of Er^{3+} in tetragonal sites in CdF_2 has been measured by Zverev *et al.*[15] at *X*-band frequencies, and the results are reviewed here for comparison. The Er concentration was ~0.1%. For $T<4.2°K$, T_1 varied inversely with temperature, where $T_1 \approx 5.5\times10^{-4}\ T^{-1}$ sec. Between 4.2 and 9°K, T_1 exhibited a rapid decrease from a T^{-1} law; lifetime broadening was observed above 16°K. Combining the last two data the authors conclude that T_1 cannot be described by a power law of the type $T_1 \propto T^{-n}$. Instead, the relaxation appears to be influenced by the presence of a nearby excited state.

Ytterbium

In Fig. 9 are shown the temperature dependences of T_1 for Yb^{3+} in cubic and tetragonal crystal fields. The resonance spectrum for this 0.17% Yb sample was very clean, consisting only of approximately equally intense cubic and tetragonal resonances. Both the $g_{\parallel}$ and $g_{\perp}$ tetragonal resonances exhibited nearly equal relaxation rates inversely proportional to T at the lowest temperatures. With increasing temperature the $g_{\perp}$ resonance T_1 decreases rapidly reaching a T^{-9} dependence by 10°K. Within experimental error, the high-temperature data may also be fitted by an exponential law with an exponent $\approx 125/T$. T_1 for the $g_{\parallel}$ resonance shows an anomalous behavior for $3.5<T<10°K$, which we do not understand at present.

The cubic resonance T_1 gradually departs from a T^{-1} law with increasing temperature. The data for $5°<T<20°K$, after subtraction of the T^{-1} dependence, can, within experimental accuracy, be fitted to either a $T^{-(3.5\pm0.5)}$ or $\exp(38/T)$ temperature dependence. The cubic Yb^{3+} resonance is unique among the rare-earth ions studied in that it is observable up to ≈115°K; the tetragonal Yb^{3+} resonances are not observable at 77°K. The temperature dependence of the cubic resonance linewidth was examined between 50–105°K. At low temperatures the peak-to-peak linewidth of the derivative of absorption is temperature-independent with a value of ≈25 G. The additional temperature-dependent linewidth $\Delta H(T)$ above 50°K is plotted in Fig. 10. If this width is attributed to lifetime broadening, its temperature dependence should reflect that of T_1 in this temperature range. The data in Fig. 10 shows a slow departure from a T^{-9} law. The temperature dependence for Raman relaxation of a Kramers ion, assuming a simple Debye lattice and characteristic temperature θ_D, is $T^9 \int_0^{\theta_D/T} [x^8 e^x dx/(e^x-1)^2]$, the T^9 limit applying only when $T<0.1\theta_D$. Since $\theta_D \approx 480°K$[16] for

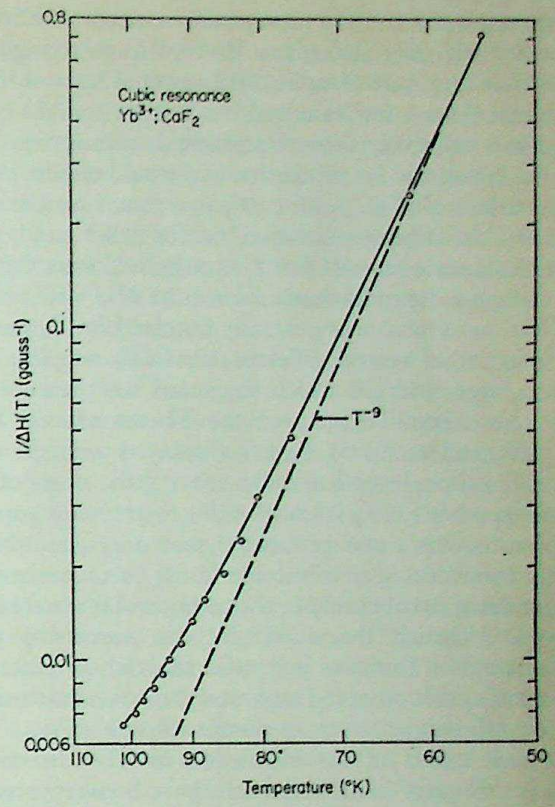


FIG. 10. Additional temperature dependence component $\Delta H(T)$ of the linewidth of Yb^{3+} in cubic sites in CaF_2.

CaF_2, the deviation from T^{-9} is expected; however, it should be more pronounced. The above integral expression with $\theta_D \sim 680°K$ provides a better fit to the data. There are several possible sources of this discrepancy; the validity of considering $\Delta H(T)$ as a simple additive contribution, the inadequacy of a simple Debye vibrational frequency spectrum for CaF_2 and the above integral expression,[17] and the possible enhanced Raman relaxation rate due to the mass excess of the rare-earth impurity.[18]

From estimates in I (see Fig. 5), the first excited state for Yb^{3+} in a cubic field in CaF_2 may be ≈600–700 cm^{-1} above the ground state. Thus T_{1R}, which is proportional to Δ^4, would be much longer than for other rare earths. This would account for the cubic-field resonance being observable at such comparatively high temperatures. Since a Δ of this size is probably greater than the highest energy branch in the phonon spectrum of CaF_2, a simple Orbach process would not be allowed. Comparing the observed broadening with Raman relaxation, a value of $M_R^2 \approx 10^5$ cm^{-2} is obtained, which is necessarily uncertain due to the lack of good agreement in fitting the temperature dependence. Extrapolating this relaxation mechanism for $T=10-20°K$ yields T_1 values several orders of magnitude longer than observed. If the higher

[15] G. M. Zverev, L. S. Kornienko, A. N. Prokhorov, and A. I. Smirnov, Fiz. Tverd. Tela 4, 392 (1962) [English transl.: Soviet Phys.—Solid State 4, 284 (1962)].

[16] D. R. Huffman and M. H. Norwood, Phys. Rev. 117, 709 (1960).

[17] M. J. Weber, Phys. Rev. 130, 1 (1963).

[18] J. G. Castle, Jr., D. W. Feldman, and P. G. Klemens, *Advances in Quantum Electronics* (Columbia University Press, New York, 1961), p. 414.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

temperature T_1's for the tetragonal resonances are fitted to a T^{-9} law, Eq. (15) gives $M_R^2 \approx 30\Delta^2$. A tetragonal field is not expected to split the Γ_8 excited state of $^2F_{7/2}$ by more than a few hundred cm^{-1}. This would again suggest a relatively large first excited-state separation for the tetragonal ion relaxation and would require very large values of M_R^2, $\lesssim 10^6$ cm^{-2}, to account for the observed T_1's. Orbach relaxation, on the other hand, via an excited state at ≈ 100 cm^{-1}, as suggested from the T_1 temperature dependence, would require $M_O^2 \approx 10^5$ cm^{-2}.

Good agreement was generally obtained for T_1 measurements from several different Yb:CaF_2 samples. No effects were observed which suggested the presence of any cross-relaxation between the Yb resonances. For $H \| [100]$, all transient signals displayed a single exponential time dependence. By varying the angle of H with respect to [100], it was possible to gradually superimpose the cubic and tetragonal resonances, at which point a short decay attributable to the cubic line and a longer decay attributable to the tetragonal line were observed. Although the uncertainty in measuring the decay times in this case is greater than when either of the transients is observed separately, the measurements are of sufficient accuracy to permit identification.

Measurements of the relaxation of Yb^{3+} in cubic sites in CdF_2 at X-band frequencies have been reported.[19] At liquid helium temperatures the relaxation of a 0.1% Yb sample is described by $T_1 = (6.7 \pm 0.7) \times 10^{-4}\, T^{-1}$ sec. No resonances were observed at liquid hydrogen temperatures. More recent investigations[20] at 3.2 and 9.1 kMc/sec in samples containing 0.01, 0.05, and 0.1% Yb demonstrate that resonances are observable at 57°K. The low-temperature relaxation times are concentration-dependent as we have noted for Ce^{3+} and show an unexpected and unexplained dependence upon resonance frequency, $T_1 \propto \delta$. From g shift, intensity and optical measurements, an excited state at $\simeq$50–70 cm^{-1} and an over-all $^2F_{7/2}$ splitting of ≈ 116 cm^{-1} is inferred. These values are much smaller than those estimated for Yb:CaF_2.

DISCUSSION

The magnitude and temperature dependence of T_1's for Kramers ions at higher temperatures, where T_1 deviates from a T^{-1} law, have been compared with theoretical expressions for two-phonon relaxation processes in Eqs. (13)–(16). A single excited state was considered with an energy Δ obtained either from the exponential temperature dependence for Orbach relaxation or from estimates in I. The experimental results indicate that for Orbach and Raman relaxation, the effective sums of orbit-lattice matrix elements M_O^2 and M_R^2 are typically in the order of 10^4–10^5 cm^{-2}. These values are comparible with those obtained using the static crystal-field parameters in I and the approximations discussed earlier relating them to the strength of the orbit-lattice coupling. The accuracy of the determination of M^2 is dependent upon estimates made in arriving at Eqs. (10), (13), and (15). For example, since the average phonon velocity v enters to the fifth and tenth powers, respectively, in the Orbach and Raman T_1's, a small error in the choice of an effective v can greatly affect the magnitude of the predicted T_1. Because of this and the uncertainty in Δ, only the magnitude of M^2 is significant.

TABLE I. Comparison of low-temperature relaxation times and impurity concentration for several rare earths in tetragonal sites in CaF_2. The rare-earth concentration is the total concentration, not the number in the particular axial site.

Ion	Concentration (%)	T_1T (sec °K)
Ce^{3+}	0.08	3.2×10^{-3}
Yb^{3+}	0.17	1.2×10^{-3}
Ce^{3+}	0.18	1.7×10^{-3}
Nd^{3+}	0.28	7.4×10^{-4}
Ce^{3+}	0.34	3.1×10^{-4}
Tb^{3+}	0.5	5.5×10^{-5}
Ce^{3+}	1.6	4×10^{-5} (2–4°K)

The above value of M^2 should be useful in estimating the magnitude and dominant relaxation process for other rare earths in CaF_2 when the excited-state energies are known. It should be noted, however, that some apparent discrepancies exist. M_O^2 for the tetragonal (I) and trigonal (II) Er resonances are very different and are beyond the extremes of the above range of M^2 values. The large M_O^2 for the trigonal resonance may arise from the strong crystalline fields associated with the O^{2-} substituted at a nearest-neighbor F^- site; however, Er was the only ion in a trigonal site whose relaxation was studied, and hence this conjecture could not be verified. The relaxation of Yb in cubic sites at temperatures 10–20°K is inconsistent with the magnitude and temperature dependence of normal two-phonon processes, if our picture of the energy levels is correct. Enhanced relaxation rates having different temperature dependences arising from the presence of lattice defects and paramagnetic empurities have been discussed[12]; however, for a Kramers salt a T^{-5} dependence is expected which is beyond the estimated error of the temperature dependence for Yb^{3+}. The relaxation may be associated with the same mechanism responsible for the low-temperature relaxation discussed below.

When Eqs. (10)–(16) are compared with measured T_1's for Kramers ions, it is immediately apparent that the present theory is insufficient to account for the fast relaxation times observed in the low-temperature region, $T \lesssim 4$°K, where T_1 is usually proportional to T^{-1}. As an estimate, consider a Zeeman admixing coefficient $(2\beta\Lambda/\Delta)\langle a|\mathbf{H}\cdot\mathbf{J}|c\rangle \approx 10^{-2}$ and sums $\sum|\langle a|V_n{}^m|c\rangle|^2 \approx 10^5$ cm^{-2}. Substitution into Eqs. (10) and (12) yields

[19] V. K. Konyukhov, P. P. Pashimin, and A. M. Prokhorov, Fiz. Tverd. Tela 4, 246 (1962) [English transl.: Soviet Phys.—Solid State 4, 175 (1962)].

[20] P. P. Pashinin and A. M. Prokhorov, *Paramagnetic Resonance* (Academic Press Inc., New York, 1963), p. 197.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$T_{1D} \approx 0.1$ sec for Kramers ions at 2°K compared to observed values of ≈ 1 msec. If one assumes that the effective summations of dynamic crystal field matrix elements in direct, Orbach, and Raman relaxation have approximately equal values, i.e., M_D^2 (non-Kramers) $\sim M_O^2 \sim M_R^2$, then the predicted relative magnitudes of the one- and two-phonon relaxation processes are incompatible with experiment. In the case of Ce^{3+}, the value of T_1 at low temperatures is concentration-dependent. A tabulation of T_1T for other rare earths in CaF_2 is given in Table I and shows a correlation with rare-earth concentration. (Concentration-dependent T_1's have been observed in other crystals containing rare-earth and iron transition group elements.[21]) Finally, equally fast low-temperature T_1's have been reported for trivalent rare earths in CdF_2[15,19] and for divalent rare earths in CaF_2.[22,23] Therefore, although the relaxation times at low temperatures generally exhibit a T^{-1} dependence characteristic of one-phonon processes, their magnitude and concentration dependence indicate that they are not derived from normal direct relaxation arising from orbit-lattice interactions.

An adequate explanation of the fast low-temperature relaxation rates of trivalent rare-earth ions in CaF_2 must account for the T^{-1} dependent of T_1, the concentration-dependent effects, the variation of T_1 with angle of the applied magnetic field, and the absence of any apparent phonon bottleneck. Also, the fast relaxation has been observed for Kramers ions in lattice sites having cubic, tetragonal, and trigonal crystal-field symmetries.

The possibility exists that some divalent rare-earth ions may be introduced into CaF_2 along with the trivalent ions. If they are non-Kramers ions, they should relax quickly and could possibly influence the relaxation of trivalent Kramers ions via cross-relaxation. This explanation is insufficient for several ions studied, however, since the accompanying non-Kramers ions have nonmagnetic ground states, an example is Yb^{2+} which is isoelectric with diamagnetic $Lu^{3+}(^1S_0)$. Divalent ion resonances were not readily apparent in the resonance spectrum. (Cross-relaxation could occur to excited state doublets, however, depending upon Δ, they would be depopulated at low temperatures which would reduce their effectiveness.) Finally, the orbit-lattice coupling for divalent ions would probably be less than for equivalent trivalent ions due to the smaller effective nuclear charge.

Enhanced relaxation rates may be caused by local vibration amplitude variations associated with the introduction of the paramagnetic impurity. In our case, this lattice defect has both a mass and charge difference and, for axial sites, has related substitutional or interstitial impurity defects. The relaxation would exhibit a T^{-1} temperature dependence and be a function of impurity concentration in certain ranges. Although some studies have been made of relaxation associated with defect sites,[12,24,25] detailed calculations applied to specific examples are needed to establish the importance of this mechanism at low temperatures where one-phonon processes are dominant.

Multiexponential transient recoveries were frequently observed at low temperatures in the T^{-1} region. We shall consider two cases where such behavior may be expected, cross-relaxation within an inhomogeneous spin system and cross-relaxation to another faster relaxing spin system. If a short, intense rf pulse applied to an inhomogeneously broadened spin system excites only a portion of it, then a decay time associated with the attainment of internal equilibrium within the total spin system may exist in addition to the time required to attain equilibrium with the lattice. The linewidths of rare earths in CaF_2 are usually larger than calculated dipolar widths and are believed to be inhomogeneously broadened. Several ions show transient recoveries from rf saturation having an initial fast temperature-independent but concentration-dependent exponential component. As the rf excitation pulse is lengthened, the initial recovery component became less prominent relative to the longer temperature-dependent component ascribed to spin-lattice relaxation. For Ce^{3+} at low temperatures, the initial component has characteristic times of ≈ 300, 130, and 40 μsec, respectively, for Ce concentrations of 0.18, 0.3, and 1.6%. This dependence is qualitatively as expected from the cross-relaxation theory of Bloembergen *et al.*[26]

If some paramagnetic species having a faster spin-lattice relaxation time is present, rare-earth ions in CaF_2 may lose their energy more rapidly via cross-relaxation to the other system than by relaxing directly to the lattice. Depending upon the nature of the fast relaxing system and the relative spin-lattice and cross-relaxation rates, an observed T_1 having a variety of concentration and temperature dependences can be obtained.[27–30] Rannestad and Wagner[30] have discussed some simple rate equations and time-dependent solutions pertinent to the expected behavior of such coupled spin systems. The reader is referred to their article,

[21] See, for example, papers in *Quantum Electronics* (Columbia University Press, New York, 1960); *Advances in Quantum Electronics* (Columbia University Press, New York, 1961); S. A. Peskovatskii and A. N. Chernets, Fiz. Tverd. Tela 4, 665 (1962) [English transl.: Soviet Phys.—Solid State 4, 484 (1962)]; A. A. Manenkov and A. M. Prokhorov, Zh. Eksperim. i Teor. Fiz. 38, 729 (1960) [English transl.: Soviet Phys.—JETP 11, 527 (1960)] and Ref. 29.

[22] E. S. Sabisky and H. R. Lewis, Proc. IEEE 51, 53 (1963); H. R. Lewis and E. S. Sabisky, Phys. Rev. 130, 1370 (1963).

[23] C. Y. Huang (private communication).

[24] B. I. Kochelaev, Dokl. Akad. Nauk SSSR 131, 1053 (1960) [English transl.: Soviet Phys.—Doklady 5, 349 (1961)].

[25] P. G. Klemens, Phys. Rev. 125, 1795 (1962).

[26] N. Bloembergen, S. Shapiro, P. S. Pershan, and J. O. Artman, Phys. Rev. 114, 445 (1959).

[27] J. H. Van Vleck, *Advances in Quantum Electronics* (Columbia University Press, New York, 1961), p. 388.

[28] J. C. Gill and R. J. Elliott, *Advances in Quantum Electronics* (Columbia University Press, New York, 1961), p. 399.

[29] N. Bloembergen and P. S. Pershan, *Advances in Quantum Electronics* (Columbia University Press, New York, 1961), p. 373.

[30] A. Rannestad and P. E. Wagner, Phys. Rev. 131, 1953 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

pp. 1958–1960, for an example of the type and range of arguments that can be presented to explain T_1 results. Their general extension to rare-earth-ion relaxation in CaF_2 is straightforward and will not be repeated here.

Although qualitative agreement might be obtained for our CaF_2 results using the above approach and appropriate models, it would be less than convincing without some knowledge regarding the origin and nature of the fast-relaxing species. The presence of other paramagnetic impurities, particularly fast-relaxing iron transition-group elements, could be important for relaxation. These impurities, however are definitely present in concentrations of less than one part in 10^4 by weight as determined by spectrochemical analysis. No distinct resonances characteristic of such ions were generally observed, although they may have been broad and hence escaped detection. Pairs, possibly exchange coupled, or clusters of rare-earth ions may be present and may account for some of the additional spectral lines observed in CaF_2 at higher rare-earth concentrations. Relaxation transitions between relatively near non-time-conjugate states associated with such clusters could be very rapid. Finally, fast relaxing centers may also be associated with the zero-field resonances noted[5,31] in Ce:CaF_2.

If the temperature-dependent relaxation times at low temperatures are characteristic of some fast-relaxing species, then the cross-relaxation time must be very fast and not the over-all relaxation bottleneck. The results for Ce^{3+}, however, show that for 1.6% Ce, an approximately temperature-independent time of ≈ 45 μsec is observed at 2–4°K suggestive of a cross-relaxation bottleneck. When the spin-lattice and cross-relaxation rates have comparable magnitudes, the relaxation of the coupled spin systems is expected to exhibit a multiexponential decay in which the relative importance of each term and its characteristic decay rate will be a function of the lattice temperature.

Until more is known about the numerous, concentration dependent lines in the EPR and optical spectra of rare-earth ions in CaF_2 and the importance of other paramagnetic impurities, the above explanation of the fast low-temperature relaxation times via cross-relaxation to other spin systems is only speculative, albeit a reasonable possibility at present.

[31] D. E. Kaplan, Bull. Am. Phys. Soc. 8, 468 (1963).

CONCLUSIONS

We have surveyed the temperature dependence of spin-lattice relaxation of several trivalent rare earths in CaF_2. The relaxation times of Kramers ions at low temperatures, which usually vary as T^{-1} and exhibit a concentration dependence, are about two orders of magnitude faster than predicted for orbit-lattice coupling. The nature of the operative relaxation mechanism which explains these results remains to be established. More intensive, detailed study of selected examples would be helpful. The magnitude and temperature dependence of two-phonon relaxation processes are generally in reasonable order-of-magnitude agreement with predictions of relaxation times using estimated crystal-field interactions. As expected, ions in lattice sites having different crystal-field symmetry and strength have different relaxation properties. The comparison between theory and experiment must of necessity be only approximate since current spin-lattice relaxation theory involves a simplified treatment of the lattice vibrational properties, but more important for rare earths in CaF_2, detailed knowledge of the energy levels, ion wave functions, and strength of the dynamic crystal field are lacking. The interpretation of the present results will benefit greatly as spectroscopic data on the crystal-field energy levels becomes available, at which time many uncertainties in the present comparison of theory and experiment will be removed.

ACKNOWLEDGMENTS

Arnold Fine's assistance in all phases of gathering the experimental data was very valuable, and the able efforts of Frederick G. Garabedian in the preparation and growth of the single crystals are greatly appreciated. Also we acknowledge valuable discussions with C. D. Jeffries, D. E. Kaplan, Z. Kiss, W. Low, P. Nutter, and L. Rimai.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Ultrasonic Absorption at Microwave Frequencies and at Low Temperatures* in MgO and Al_2O_3

I. S. CICCARELLO† AND K. DRANSFELD
Department of Physics, University of California, Berkeley, California
(Received 22 January 1964)

Measurements are presented of the ultrasonic absorption in MgO and Al_2O_3 for transverse and longitudinal waves in the low-temperature region. In contrast to some theoretical expectations, the absorption of longitudinal waves is not negligible but even stronger than for transverse waves. In the theoretical discussion it is shown that this strong longitudinal absorption probably is caused by three-phonon processes.

INTRODUCTION

THE interaction of ultrasonic waves with thermal phonons in dielectric crystals shows a very different behavior at high and at low temperatures. At high temperatures when the ultrasonic period $2\pi/\omega$ is much larger than the lifetime θ of most termal phonons the measured absorption can be well described by a mechanism first proposed by Akhiezer.[1-3] However, in the low-temperature region, defined by ($\omega\theta \gg 1$) experimental difficulties have prevented until recently a careful study of the hypersonic absorption, although a theory for the absorption process was already published by Landau and Rumer[4] more than 25 years ago. They pointed out that through the anharmonic cubic terms in the elastic lattice energy *transverse* acoustic waves or phonons can interact with thermal phonons in processes involving three phonons. For the absorption of *longitudinal* acoustic waves these three-phonon processes were considered unimportant on the basis of energy and momentum conservation as will be discussed later. A finite absorption was thought to result only when according to Pomeranchuk[5] fourth-order terms in the elastic energy were taken into account (so-called four-phonon processes), or when according to Herring[6,7] the elastic anisotropy of the crystal was properly considered (Herring mechanism).

In the present paper we present measurements of the absorption of transverse and longitudinal ultrasonic waves in magnesium oxide and sapphire in the low-temperature region ($\omega\theta > 1$). The interesting result is that not only the transverse but also the longitudinal absorption follows the pattern expected for a three-phonon process. The magnitude of the longitudinal absorption as well as its dependence on frequency and temperature are not compatible with a four-phonon process or the Herring mechanism mentioned above.

By these results we are led to believe that longitudinal ultrasonic waves also are able to interact strongly with thermal phonons in three-phonon processes. We have already shown in a classical calculation,[8] referred to here as paper I, that these "collinear" processes lead to a strong ultrasonic absorption, and the same result will be derived here by a quantum-mechanical argument.[8a]

This strong absorption of longitudinal waves in three-phonon processes also solves the problem of a diverging thermal conductivity which was discussed in the literature[9] under the assumption that longitudinal waves cannot be absorbed in three-phonon processes. Since they are, however, absorbed, their contribution to the heat conductivity remains finite, and no problem arises in this context.

EXPERIMENTS AND RESULTS

We have measured the attenuation of sound waves in magnesium oxide and ruby at the two frequencies of 500 and 3000 Mc/sec between 4.2 and about 100°K. The samples were single crystals of rectangular shape, 4×4×25 mm, with the 4×4-mm faces polished optically flat and parallel within a few seconds. The rod axes were oriented within one degree, parallel to the [100] axis of MgO and to the c axis of ruby. The MgO samples were colorless and quoted[10] as pure while the ruby rods contained up to 0.01% of Cr_2O_3. Care was taken to avoid resonant absorption of sound by the paramagnetic ions.[11]

For the generation and detection of the ultrasonic waves we used two methods: At 500 and at 3000 Mc/sec two x-cut quartz disks with a fundamental resonance frequency of 10 Mc/sec were bonded to the sample and excited in a high harmonic. For some experiments at 3000 Mc/sec both sides of our samples were plated with films of permalloy, 90% Ni, 10% Fe,

* This work was supported by the U. S. Atomic Energy Commission.
† NATO Fellow on leave of absence from Physics Department, University of Palermo, Palermo, Italy.
[1] A. Akiezer, J. Phys. (USSR) 1, 277 (1939).
[2] H. Bommel and K. Dransfeld, Phys. Rev. 117, 1245 (1960).
[3] T. O. Woodruff and H. Ehrenreich, Phys. Rev. 123, 1553 (1961).
[4] L. Landau and G. Rumer, Z. Physik. Sowjetuuion 11, 18 (1937).
[5] I. J. Pomeranchuk, J. Phys. (USSR) 4, 259, 529, (1941); 6, 237 (1942).
[6] C. Herring, Phys. Rev. 95, 954 (1954).
[7] S. Simons, Proc. Cambridge Phil. Soc. 53, 702 (1957).
[8] R. Nava, R. Arzt, I. S. Ciccarello, and K. Dransfeld, Phys. Rev. 134, A581 (1964).
[8a] *Note added in proof.* Arguments for this mutual interaction of parallel phonons have also recently been given by S. Simons [Proc. Phys. Soc. (London) 82, 401 (1963)].
[9] P. G. Klemens, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1956), Vol. 14. p. 1.
[10] Supplied by R. Meller Co., Providence, Rhode Island.
[11] E. B. Tucker, Phys. Rev. Letters 6, 183 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

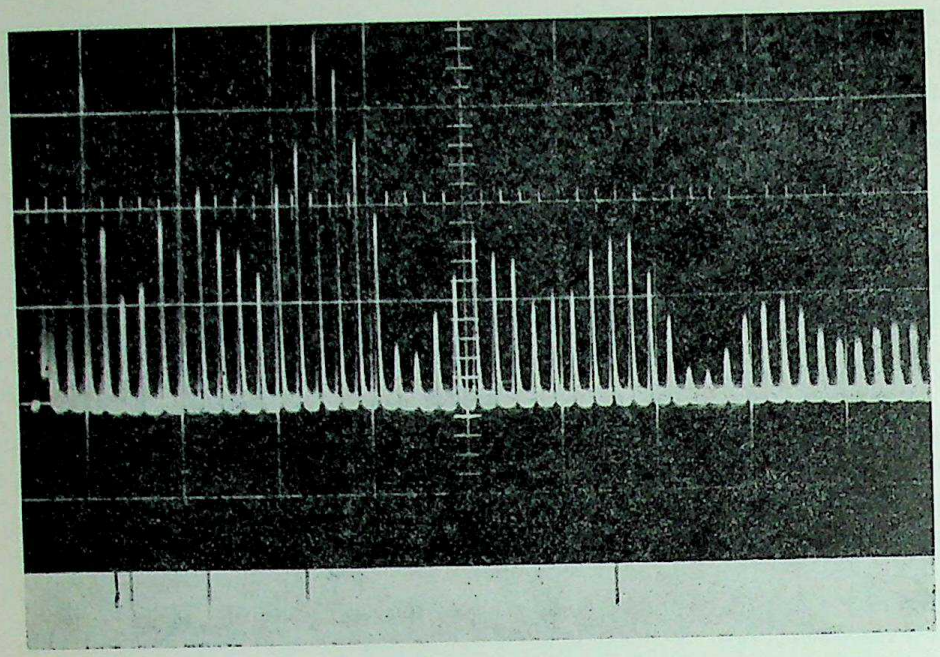


FIG. 1. Pulse pattern of 3-kMc/sec longitudinal waves in ruby at 4.2°K. The waves are excited and detected ferromagnetic resonance in thin films of permalloy (90% Ni-10% Fe) on each side of the ruby rod.

about 5000 Å thick for the generation and detection of sound by ferromagnetic resonance.[12] By using standard pulse methods sometimes more than 100 acoustic echo pulses were received.

Figure 1 shows a typical pulse pattern, the time between two consecutive pulses corresponds to the time for one round trip in the rod. The decay of the pulse pattern is hardly ever exponential. The reasons for the irregular amplitudes of the various pulses are several: the end faces may not be sufficiently parallel, the rod may not be elastically homogeneous, etc., all of these effects being temperature-independent. The absorption at a temperature T was determined from the decrease of an echo pulse relative to its magnitude at 4.2°K.

With ferromagnetic excitation we occasionally observed a drastic change in the relative amplitude of the peaks on varying the intensity or direction of the magnetic field. We believe that this is caused by interference effects which arise from mechanical or magnetic nonuniformities of the films, or the excitation of Walker modes. By improving our deposition methods and by using films of small cross section these phenomena could be avoided.

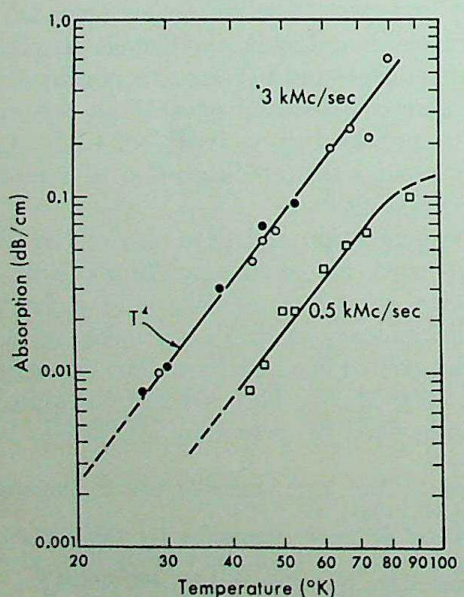


FIG. 2. Attenuation of longitudinal waves in ruby as a function of temperature.

Through our long ruby rod only longitudinal waves could be transmitted well. The propagation of transverse waves proved to be considerably more difficult. In general, only one transverse echo could be seen. This difficulty probably arise from the acoustical birefringence, so that along the c axis the acoustic wave vector and Poynting vector are not parallel, and the acoustic energy does not travel parallel to the rod axis. The absorption of longitudinal waves versus temperature is plotted in Fig. 2. Several runs at 3000 Mc/sec, with different ruby rods carrying different magnetic films gave the reproducible results of Fig. 2: The absorption increases with temperature approximately as T^4. For the investigation of the frequency dependence, measurements at 500 Mc/sec, which agree with the results of Fitzgerald and Truell,[13] are also included in Fig. 2; the frequency dependence is close to linear.

The absorption of longitudinal and transverse waves in MgO is plotted versus temperature in Figs. 3(a) and 3(b). The data for longitudinal waves, although not as accurate as the transverse data, clearly indicate that the frequency dependence is again close to linear, and the temperature dependence not far from T^4.

With longitudinal waves the acoustic power level was kept low enough to avoid the nonlinear effects reported by Shiren,[14] and for transverse waves such a power-dependent absorption could not be observed even

[12] H. Bommel and K. Dransfeld, Phys. Rev. Letters 3, 83 (1959).

[13] T. M. Fitzgerald and R. Truell, research report, Brown University, 1963 (unpublished).

[14] N. S. Shiren, Phys. Rev. Letters 11, 3 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

at considerably higher power levels than those used by Shiren. To our knowledge, for shear waves in the kMc/sec range, these are the first data on the temperature dependence in a material other than quartz.

Both for transverse and longitudinal waves in MgO a noticeable additional absorption occurred around 15°K, when the absorption was measured in zero magnetic field. Figure 3(a) also shows this extra absorption at 3000 Mc/sec, which could be removed by applying a magnetic field of the order of 1000 Oe. A similar phenomenon was observed by Bolef *et al.*[15] at lower frequency and lower temperature in a more drastic way in heavily doped MgO. More experiments are planned for a detailed investigation of this anomaly which is probably caused by paramagnetic impurities.

DISCUSSION

Our experiment results may be summarized by saying that at low temperatures for both transverse and longitudinal waves the absorption increases linearly with frequency and with the fourth power of the temperature. In perfect crystals the absorption is due only to scattering by other thermal phonons, and is possible only if the potential energy developed in powers of the strain contains cubic or higher order terms. Landau and Rumer[4] considering the cubic terms only calculated the absorption of transverse waves for low temperatures, when $\omega\theta \gg 1$. They found

$$\alpha_t \approx \tfrac{1}{4}(2Q+6R)^2(1-V_t/V_L)^2 \frac{k^4}{\rho^3 V_L{}^{10}\hbar^3}\omega T^4, \qquad (1)$$

where Q and R are constants of the order of 5×10^{12} for ruby and MgO, k is the Boltzmann constant, and ρ the density. The agreement with our results on MgO is satisfactory. The absorption is of the right order of magnitude and varies with temperature as T^4, as expected. Expression (1) was derived only for *transverse* waves, with the observation that the conservation of energy and momentum rules out three-phonon processes for the absorption of *longitudinal* waves.

Pomeranchuk[5] invoked the much smaller fourth-order terms of the elastic energy to account for the scattering of longitudinal phonons. The absorption according to his calculations should be proportional to ω^2T^7, which is in strong disagreement with our experiments, both for ruby and MgO. In addition, Orbach[16] has already estimated, for the case of quartz, that Pomeranchuk's calculations give an absorption which is several orders of magnitude smaller than the observations. We may therefore safely conclude that a four-phonon process is not responsible for the observed absorption of longitudinal waves.

On the other hand, the Herring mechanism[6] also cannot account for our results. The absorption due to this process was calculated by Simons[7] for cubic crystals. For a longitudinal wave propagating parallel to the (100) axis he found

$$\alpha = A\omega^2T^3. \qquad (2)$$

The quadratic frequency dependence is not compatible with our observations. The factor A depends on the crystal and can be estimated for MgO, with the result that for our frequencies and temperatures the absorption according to (2) is about three orders of magnitude smaller than the measured values. For ruby the absorption by the Herring mechanism has not been calculated numerically, but should be proportional to ω^3T^2, which agrees even less with our experiments.

In view of the fact that our data with longitudinal waves behave as if they are caused by three-phonon processes, and in view of the difficulty in explaining them in any other way, we proceed to re-examine in more detail the absorption of longitudinal waves by three-phonon processes. We have already shown classically in paper I that in isotropic solids the absorption of longitudinal waves by three-phonon processes is of similar magnitude as the absorption of transverse waves. But the fact that a longitudinal acoustic wave can interact in three-phonon processes with almost as many thermal phonons as a transverse wave may be shown in a simpler argument, which will be presented here:

The conservation of energy and momentum requires

$$\omega_1+\omega_2=\omega_3, \qquad (3)$$

$$\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}_3, \qquad (4)$$

where the index "1" refers to the ultrasonic phonon. Since in our temperature range $|k_1| \ll |k_{2,3}|$ the momentum conservation (4) can be expressed in terms of the angle α between the acoustic wave vector $\mathbf{k}_1$ and $\mathbf{k}_2$ (see Fig. 4).

$$\cos\alpha = (|k_3| - |k_2|)/|k_1|,$$

or, because of

$$|k_1| = \omega_1/v_1, \quad |k_{2,3}| = \omega_{2,3}/v_{\rm ph},$$

we can write for (4)

$$\cos\alpha = \frac{v_1}{v_{\rm ph}}\left(\frac{\omega_3-\omega_2}{\omega_1}\right), \qquad (5)$$

where α is the angle under which a thermal phonon with velocity $v_{\rm ph}$ can interact with an acoustic wave of velocity v_1 with conservation of energy and momentum.

But if energy is not exactly conserved because of the uncertainty in the energy arising from a finite relaxation time of θ of the thermal phonons, we find

$$\omega_3-\omega_2-\omega_1 \leqq 1/\theta$$

and $\cos\alpha$, according to (5), can assume any value between the two limits

$$\cos\alpha_{\max} = \frac{v_1}{v_{\rm ph}}\left(1-\frac{1}{\omega_1\theta}\right); \quad \cos\alpha_{\min} = \frac{v_1}{v_{\rm ph}}\left(1+\frac{1}{\omega_1\theta}\right). \qquad (6)$$

[15] D. I. Bolef, J. DeKlerk, and R. B. Gosser, Rev. Sci. Instr. 33, 631 (1962).

[16] R. L. Orbach, thesis, University of California, Berkeley, 1960 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

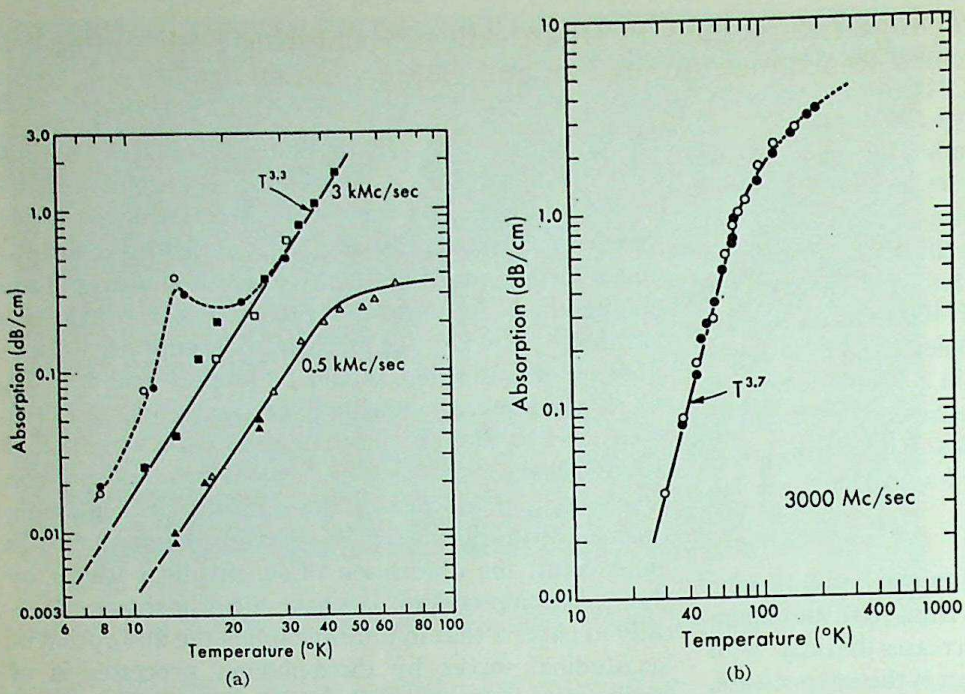


FIG. 3. Attenuation of ultrasonic waves in MgO as a function of temperature. The dashed curve gives the attenuation in zero external magnetic field; the continuous ones in a field of about 4500 Oe. (a) Longitudinal waves, (b) transverse waves.

Clearly, because of the uncertainty in energy, a thermal phonon can interact with an acoustic phonon between the angles $\alpha_{\max}$ and $\alpha_{\min}$ without violating the conservation of energy and momentum. For *transverse* waves, for example, we find from (6)

$$\alpha_{\max}-\alpha_{\min}=\Delta\alpha=\frac{2\cos\alpha}{\sin\alpha}\frac{1}{\omega_1\theta}. \tag{7}$$

This corresponds to a relative solid angle

$$\frac{\Delta N}{N}=\frac{2\pi\sin\alpha}{4\pi}\frac{2\cos\alpha}{\sin\alpha}\frac{1}{\omega_1\theta}=(\cos\alpha)\frac{1}{\omega_1\theta}. \tag{8}$$

For *longitudinal* waves ($v_1=v_{\rm ph}$, $\alpha_{\min}=0$, collinear case) we find from (6)

$$\tfrac{1}{2}(\Delta\alpha)^2=1/\omega_1\theta,$$

and for the relative solid angle

$$\frac{\Delta N}{N}=\frac{\pi\alpha_{\max}^2}{4\pi}=\frac{1}{2}\left(\frac{1}{\omega_1\theta}\right). \tag{9}$$

On comparing (8) and (9) it is evident that the two solid angles are almost identical for any relaxation time θ, as long as $\omega_1\theta>1$, and the dispersion is neglected.

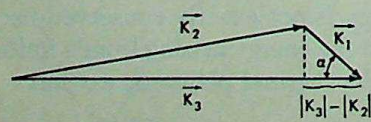


FIG. 4. Momentum conservation for the absorption of an acoustic phonon with wave vector $\mathbf{k}_1$ by a thermal phonon of wave vector $\mathbf{k}_2$. $\mathbf{k}_3$ is the wave vector of the final phonon in this three-phonon process.

One should therefore expect that a quantum-mechanical treatment of the ultrasonic absorption due to three-phonon processes should give very similar results for transverse and longitudinal waves. We made such a calculation, following the method of Landau and Rumer,[4] and we found for the absorption of longitudinal waves due to the interaction with collinear thermal phonons:

$$\alpha_L\approx\tfrac{1}{2}[6P]^2\frac{k^4}{\rho^3V_L{}^{10}\hbar^3}\omega T^4. \tag{10}$$

P is a constant of the same order of magnitude as Q and R in Eq. (1), and therefore the absorption of longitudinal waves is not too different from the absorption of transverse waves.

In his argument we have neglected the dispersion. As shown in paper I this neglect is justified as long as

$$\omega\theta\leqq 2\pi(T_D/T)^2,$$

with T_D being the Debye temperature of the crystal. For our samples below 60°K $2\pi(T_D/T)^2\approx 1500$ and for our frequencies $\omega\theta$ is indeed smaller than 1500, if the mean free path of the thermal phonons remains below 1 mm. Thus, the dispersion can be neglected.

It seems to be a characteristic feature of cubic crystals that the absorption of longitudinal waves is noticeably stronger than for transverse waves. This is evident from our measurements on MgO and was also found in Ge by Dobbs *et al.*[17]

[17] E. R. Dobbs, B. B. Chick, and R. Truell, Phys. Rev. Letters 3, 332 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Role of *s-d* Exchange Interactions in Dilute Alloys Exhibiting Both Low-Temperature Resistance Anomalies and Superconductivity*

E. W. COLLINGS, F. T. HEDGCOCK, AND Y. MUTO†
The Franklin Institute Laboratories for Research and Development, Philadelphia, Pennsylvania
(Received 28 October 1963; revised manuscript received 28 February 1964)

Measurements of the magnetic susceptibility of some dilute Zn-Mn alloys are reported in the temperature range where this alloy system is known to exhibit low-temperature resistance anomalies. All of the alloys studied obey a Curie law with the most concentrated alloy (0.43-at.% Mn) exhibiting deviations in the neighborhood of 6°K. The Curie constant indicates that manganese when dissolved in zinc has a spin value of $\frac{5}{2}$ which is temperature-independent between 5 and 273°K. Calculations of the exchange integral J from the magnetic properties of dilute Zn-Mn alloys have been made from the Néel temperature, the magnetoresistivity as a function of field strength and the suppression of the superconducting critical temperature based on theoretical equations which assume a scalar *s-d* interaction between the paramagnetic ions and conduction electrons. All of the derived values of J lie in the range of $(1.4\pm0.2)\times10^{-12}$ erg which would lead one to conclude that the magnetic interaction giving rise to these various effects is of the same nature. A graphical method for separating the resistance minimum from the total resistivity, in the presence of a resistance maximum produced by magnetic ordering in the Zn-Mn system, is presented.

INTRODUCTION

THEORY predicts[1,2] and some experiments substantiate[3–5] the fact that the critical temperature is suppressed in superconducting elements containing small concentrations of transition metal impurities. Recently, measurements have been reported[5] on the suppression of the superconducting transition temperature in a system that had been earlier reported to exhibit a low-temperature resistive anomaly.[6] It has been shown theoretically[7,8] and illustrated experimentally for the[9] Zn-Mn system that a sharply energy-dependent conduction-electron relaxation time in the neighborhood of the Fermi energy will produce a resistance rising with decreasing temperature to some finite value at absolute zero. It is observed in the Zn-Mn system that the resistive characteristic flattens off following the so-called resistance minimum.[9] A preliminary report[10] of the magnetic properties of this system show a weak antiferromagnetism and it is believed that although the resistivity is decreased by the decrease of spin scattering on ordering, the decrease is not sufficiently great to result in a resistive maximum. It is the purpose of the present paper to present more complete data on the spin value of manganese in zinc and to calculate and compare values for the *s-d* exchange integral J found from electron transport and magnetic properties. A comparison of the value of J found for the Zn-Mn system will be made with the values of J found for a system possessing only a resistive anomaly (Cu-Mn) and only a superconducting transition (Al-Mn).

MAGNETIC SUSCEPTIBILITY RESULTS AND DISCUSSION

Table I lists the results of the magnetic susceptibility measurements.

Nominal manganese concentration was obtained from a plot of "relative residual resistance" versus "weighed-out" concentration. Figure 1(a) shows the collected data from several series of alloys. In Fig. 1(b) the room temperature susceptibilities of some rolled strips and the cast pieces used in the present temperature runs are plotted against concentration obtained from relative residual resistances of rolled strips. The susceptibility measurements were regarded as being the most accurate and deviations of the points from the best straight line were attributed to concentration errors. Figure 1(b) was therefore used to obtain the sample concentration used in subsequent calculations.

In the temperature runs, relative susceptibility was measured and converted to susceptibility from the separately-measured room temperature values. Figure 2 shows the alloy susceptibility plotted against reciprocal temperature in the region between room temperature and liquid nitrogen. In Fig. 3 the inverse of the manganese component of susceptibility (ΔX^{-1}) is shown as a function of absolute temperature in the temperature range between liquid nitrogen and liquid helium. An insert shows the behavior of the most concentrated alloy in the vicinity of the Néel point. If one

* This work was supported through the Franklin Institute funds for self-sponsored research and partial support by Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio.

† Visiting scientist on leave from the Research Institute for Iron, Steel and other Metals, Tohoku University, Sendai, Japan.

[1] W. Baltensperger, Helv. Phys. Acta., 32, 197 (1959).
[2] A. A. Abrikosov and L. P. Gorkov, Zh. Eksperim. i Teor. Fiz. 12, 1243 (1961) [English transl: Soviet Phys.—JETP 39, 1781 (1960).
[3] G. Boato, G. Gallinaro, and I. Rizzuto, Phys. Letters 5, 20 (1963).
[4] M. Merriam and D. P. Seraphim, Colgate Superconductivity Conference (to be published).
[5] G. Boato, G. Gallinaro, and I. Rizzuto, Colgate Superconductivity Conference (to be published).
[6] Y. Muto, Y. Tawara, Y. Shibuya and T. Fukuroi, J. Phys. Soc. Japan 14, 380 (1959).
[7] J. Korringa and A. N. Gerritsen, Physica 19, 457 (1953).
[8] C. A. Domenicali, Phys. Rev. 117, 984 (1960).
[9] F. T. Hedgcock and W. B. Muir, Phys. Rev. 129, 2045 (1963).
[10] E. W. Collings, F. T. Hedgcock, and W. B. Muir, Proceedings of the Eighth International Conference on Low-Temperature Physics, London, 1962 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Magnetic susceptibility samples with the derived concentrations, Curie constants, and calculated magneton numbers.

Sample	Nominal concentration from Fig. 1(a) weight (%)	Concentration from ΔX data [Fig. 1 (b)] weight (%)	$10^6 X$ at room temperature emu/gm	$10^5\times$ Curie constant		Effective magneton number; P_{eff}	
				Room temp. to nitrogen	Nitrogen to helium	Room temp. to nitrogen	Nitrogen to helium
Pure Zn			−0.140				
VII	0.042	0.046	−0.077	1.5_5	1.6_2	3.8_5	3.9_4
							3.8_5[a]
IV	0.18	0.20	+0.121	7.5_5	1.6_4	4.0_8	4.1_0
II	0.43	0.40	+0.381	18.0	15.2	4.4_5	4.1_0

[a] The effective magneton number for this sample is also determined from the zero susceptibility value as shown in Fig. 2.

considers that $(dX^{-1}/dT)_{T_N}=0$ defines the Néel temperature, a value for T_N of 6±1°K is obtained for sample II. From the linear portions of Figs. 2 and 3, Curie constants were calculated and using concentrations obtained from Fig. 1(b) these were converted into effective magneton numbers. The desired results are shown in Table I from which it can be seen that the spin state of the manganese ion in zinc corresponds to a value of $\frac{3}{2}$ which would give an effective magneton number of 3.9.[11]

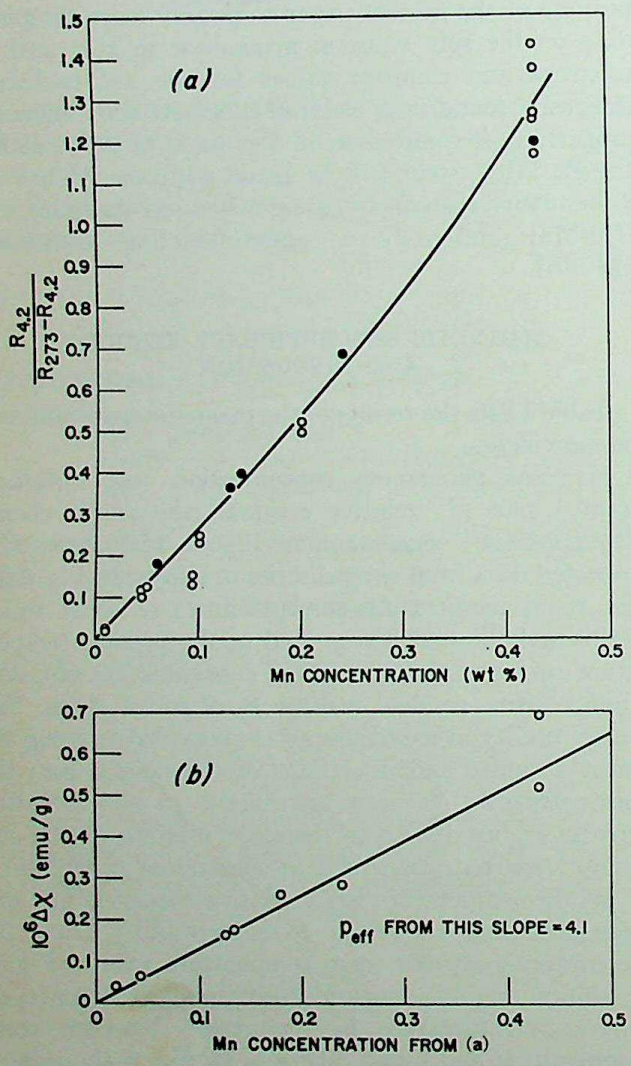


FIG. 1. (a) Resistance ratio as a function of weight percent manganese determined from spectroscopic analysis or quantity of manganese added to the melt. The solid circles represent data from the thesis of W. B. Muir, University of Ottawa. (b) The increase in susceptibility of pure zinc as a function of manganese concentration as determined from the resistance ratio of the alloys measured.

ANALYSIS OF THE MAGNETORESISTIVITY DATA FOR Zn-Mn SYSTEM

In an earlier publication[12] it was shown that the magnetoresistivity of dilute Cu-Fe alloys can be expressed as the sum of a normal magnetoresistive term ρ_n which is temperature-independent and obeys Kohlers' rule and an anomalous term ρ_s which is negative and decreases with decreasing temperature. The anomalous term ρ_s can be theoretically accounted for semiquantitatively for these alloys although the theory itself does not contain an explanation of the resistive minimum.[13,14] The magnetoresistive behavior for an alloy containing 0.11-at.% Mn is shown in Fig. 4 where it can be seen that the magnetoresistance is negative in sign and saturates at a field value of approximately 32 kOe. Shown in the insert of Fig. 4 is the magnetoresistance plotted as a function of H^2 where it can be seen that the magnetoresistance beyond the saturation value is a function of H^2. This we attribute to the normal magnetoresistive component, ρ_n. Using this device we can then separate ρ_n from ρ_{total} and obtain ρ_s as shown in Fig. 4.

Using Kasuya's expression for the term ρ_s and neglecting small terms, we can write:

$$\frac{\rho_{H\infty}}{\rho_{H_0}}=\frac{(b+C)_{H_0}}{(b+\langle j_z^2\rangle_{\text{av}})_{H_0}^2}\Big/\frac{(b+C)_{H\infty}}{(b+\langle j_z^2\rangle_{\text{av}})_{H\infty}^2}, \tag{1}$$

and

$$\rho_{H_0}=\frac{3\pi m^* c}{2h\zeta Ne^2}J^2\frac{(b+\langle j_z^2\rangle_{\text{av}})^2}{b+C}, \tag{2}$$

where

$$C=\langle j_z^2\rangle_{\text{av}}-\langle j_z\rangle_{\text{av}}\frac{2xe^x}{(e^x-1)^2},$$

[11] It should be pointed out that the earlier reported (Ref. 10) spin value of 2 for a concentrated Zn-Mn alloy resulted from a determination of the Curie constant too near the Néel temperature.

[12] Y. Muto, K. Noto, and F. T. Hedgcock, Can. J. Phys. 42, 15 (1964).

[13] K. Yosida, Phys. Rev. 107, 396 (1957).

[14] T. Kasuya, Progr. Theoret. Phys. (Kyoto) 72, 228 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

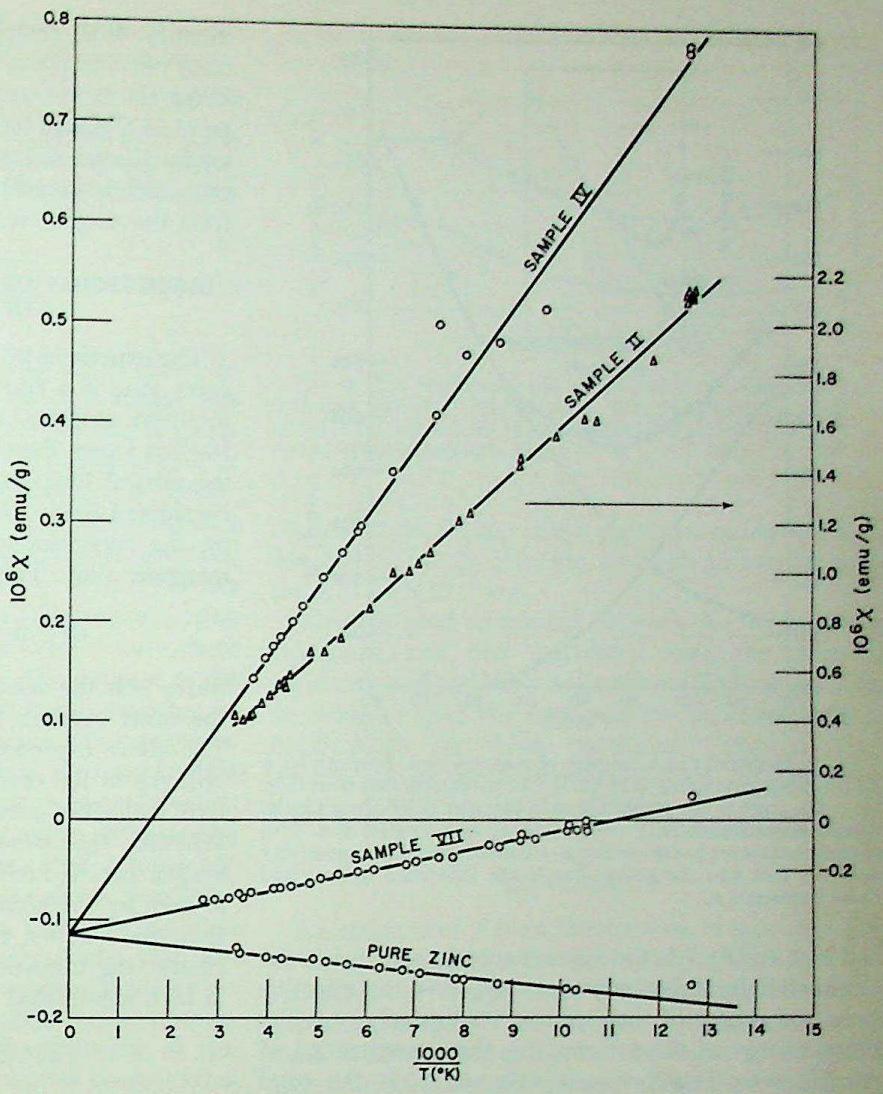


FIG. 2. Total susceptibility of the zinc manganese alloys and pure zinc as a function of reciprocal temperature between room temperature and the temperature of liquid nitrogen.

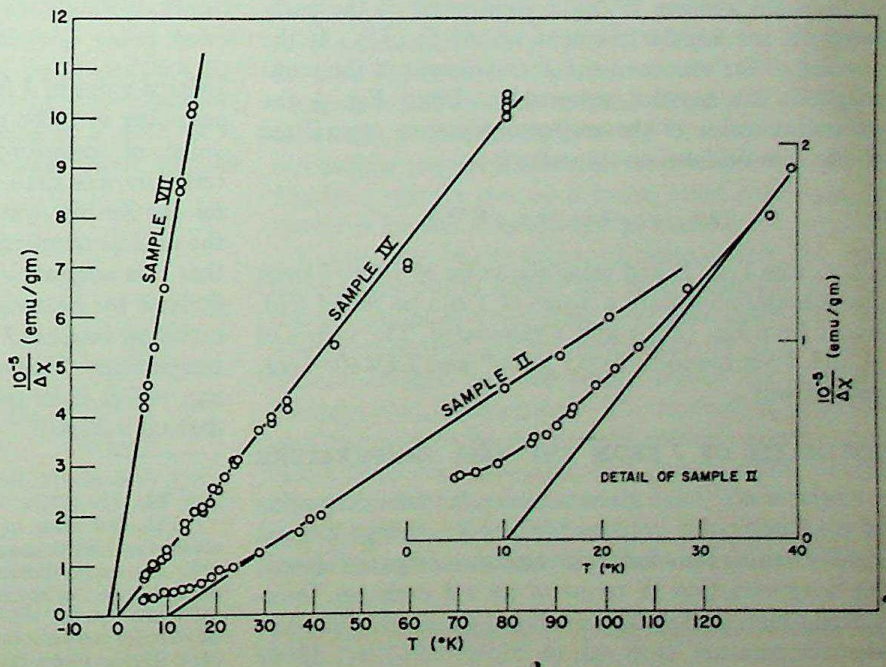


FIG. 3. The inverse of the susceptibility difference between the alloy and pure zinc as a function of temperature. The inset is an enlargement of the temperature interval below 40°K for sample II.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

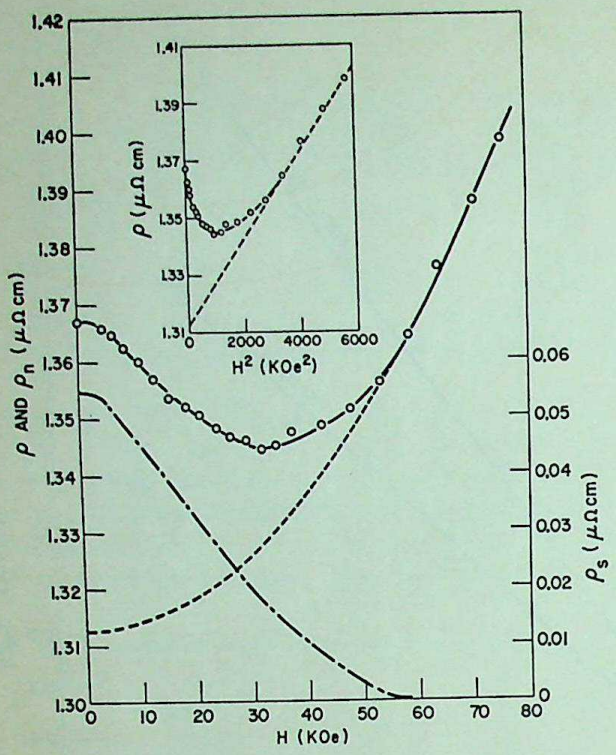


FIG. 4. Resistivity as a function of magnetic field strength for a Zn-Mn alloy containing 0.11 at.% Mn as determined from Fig. 1(a). The inset is the resistivity as a function of H^2 showing the normal magnetoresistivity component ρ_n as a dashed line. This normal component is also shown as a dashed line in the resistivity versus H plot and the dotted line is the difference $(\rho-\rho_n)$ and hence represents ρ_s.

and $b=(A/J)^2$; J is the s-d exchange integral; A is the Coulomb interaction constant; m^*, e is the effective mass and charge of the electron, respectively; ζ is the Fermi energy of the solvent; c is the concentration of paramagnetic ions per unit volume; N is the total number of atoms per unit volume; x is $g\mu_B H/kT$; $\langle j_z\rangle_{\rm av}$ is the average of the Z component of the paramagnetic ion angular moment vector j; $\langle j_z^2\rangle_{\rm av}$ is the average of the square of the Z component of the paramagnetic ion angular momentum. From Fig. 4 the saturation value of the magnetoresistance normalized to the zero-field resistivity value is

$$\rho_{H_{\rm sat}}/\rho_{H_0}\equiv\rho_{H_\infty}/\rho_{H_0}=0.960.$$

Then from Eq. (1) and using the value of $j=s=\frac{3}{2}$ from the susceptibility data a value of b can be found and, hence from Eq. (2), A and J calculated. The values of A and J are found to be 8.7×10^{-12} and 1.4×10^{-12} erg, respectively.

ESTIMATES OF J FROM THE NÉEL TEMPERATURE

Owen *et al.*[15] have given a molecular field derivation of the Curie point in terms of the s-d exchange integral J.[16] By writing the conduction electron and paramagnetic ion magnetization in terms of an s-d exchange interaction constant they find a nonzero magnetization when the temperature is equal to $S(S+1)J^2C/2k\zeta$. If we identify this temperature with the experimentally observed Néel point[17] we may then estimate a value of J. Using the most concentrated alloy which shows a Néel point of 6°K and using the Fermi energy of 1.64×10^{-11} erg for zinc yields a value of J equal to $(1.5\pm0.2)\times10^{-12}$ erg which is in good agreement with the value derived from the magnetoresistivity data.

SUPPRESSION OF THE CRITICAL TEMPERATURE IN DILUTE ALLOYS

The experimental situation at the present time indicates that if a transition metal impurity possesses a localized moment which corresponds closely to the free-ion value, then a suppression of the superconducting critical temperature will result. Baltenperger[1] has considered the influence of an s-d exchange interaction on the superconducting properties of a dilute paramagnetic alloy. The expression derived by him is

$$dT_c/dc=-2.49J^2s(s+1)\nu_0/kN, \qquad (3)$$

where ν_0 is the density of states at the Fermi level and the other symbols have the same meaning as before. The above expression is only valid if there is no spin ordering of the conduction electrons due to a cooperative magnetic interaction of the paramagnetic ions. Whether this condition is satisfied for the Zn-Mn system can be investigated by requiring that the field θ^n seen by the nonsuperconducting electrons is smaller than the thermal energy corresponding to the superconducting transition temperature. Baltensperger has, in fact, shown that

$$d\theta^n/dc=\nu_0H^2s(s+1)/6kN.$$

The inequality to be satisfied is then that

$$\partial\theta^n/\partial c<\partial T_c/dc.$$

Using a value of J found from the previous analysis and assuming ν_0 to be concentration-independent, yields a value of approximately 10°K/at.%. This value is certainly less than the value of 170°K/at.% reported for the Zn-Mn system in studies of the suppression of the critical temperature by Boato *et al.*[3] Having shown that this inequality is satisfied we can then use Eq. (3) derived by Baltensperger to calculate a value of the exchange integral J from the suppression of the critical temperature. This equation yields a value of 1.6×10^{-12} erg, which is in good agreement with the previously derived values.

---

[15] J. O. Owen, M. Browne, W. D. Knight and C. Kittel, Phys. Rev. 102, 150 (1956).

[16] This derivation neglects the nonuniform polarization of the conduction electrons and the resulting influence on the paramagnetic ion magnetization which undoubtedly forms a most significant aspect of any such calculation.

[17] Only for the simplest antiferromagnet will the Curie and Néel temperatures be equal. See C. Kittel *Introduction to Solid State Physics* (John Wiley & Sons, Inc., New York, 1953), p. 190.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

RESISTANCE ANOMALY AND THE s-d EXCHANGE INTEGRAL

According to Yosida[13] and Kasuya[14] it should be possible to obtain a value for the s-d exchange integral in a dilute alloy exhibiting an antiferromagnetic ordering from the change in resistance above and below the Néel temperature

$$\Delta\rho \equiv \rho_N - \rho_0 = \frac{3\pi msc}{2e^2\hbar\zeta N} J^2. \qquad (4)$$

This graphical method for evaluating $\Delta\rho$ is shown in Fig. 5(a) and in applying this method Yosida has assumed that the resistivity in the absence of a cooperative interaction would be temperature-independent. If the resistivity below the Néel temperature is the resultant of two opposing effects, viz., the resistance minimum mechanism and the cooperative interaction, the true $\Delta\rho$ would be $(\rho_1-\rho_0)$ of Fig. 5(a). (For the Cu-Mn alloy referred to in Yosida's theoretical paper, application of this idea would lead to more than twice his value of $\Delta\rho$ and consequently a 50% increase in the calculated value of J to 3.0×10^{-12} erg resulting in better agreement with the theoretically free manganese ion value of 3.5×10^{-12} erg.) The true situation for this model, however, is not clear when the region over which the cooperative interaction sets in severely overlaps the region in which one would otherwise expect a resistance minimum resulting in the resistivity characteristic of Fig. 5(b). In estimating $\Delta\rho$ for a Zn-Mn alloy it is better to select an alloy sufficiently dilute that the antiferromagnetic transition is more clearly separated from the minimum. As indicated earlier we are proposing that in the Zn-Mn system the influence of the cooperative magnetic interaction exactly compensates the increase in resistance with lowering temperature characteristic of the resistivity minimum alone, resulting in a flattening-off of the resistivity curve below T_N [Fig. 5(b)]. In order to obtain $\Delta\rho$ we produce the resistivity curve as a circular arc having the same radius as the solid curve above T_N resulting in a value of ρ_1 from which $\Delta\rho=\rho_1-\rho_0$ can be obtained. Applying this

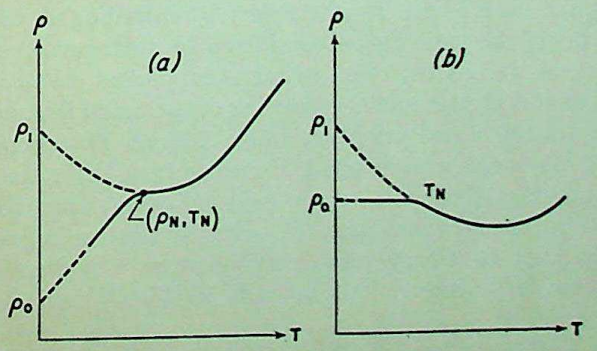


FIG. 5. (a) The resistivity as a function of temperature for a typical high concentration Cu-Mn alloy. (b) The resistivity as a function of temperature for a typical high concentration Zn-Mn alloy. The dashed lines represent the extrapolated curves.

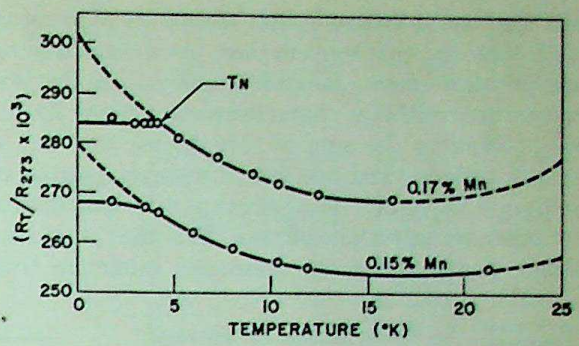


FIG. 6. The resistance ratio as a function of temperature of a 0.15 and 0.17-at.% Zn-Mn alloy. The resistivities of these two alloys is 6.7 and 6.9 $\mu\Omega$-cm, respectively. T_N is the corresponding temperature indicated in Fig. 5.

method to the Zn-Mn alloys of concentration 0.15-at.% Mn and 0.17-at.% Mn (Fig. 6) yields an average value for J of 1.5×10^{-12} erg.

There is good agreement between the value of J so calculated and that calculated from the magnetoresistance and the Néel temperature. This can be taken as evidence that the extrapolation procedure and its implications are valid, particularly since a direct application of the unmodified Yosida method would yield a $\Delta\rho$ of 0 for Zn-Mn alloys.

CONCLUSION

An estimate of J from the decrease in spin scattering on magnetic ordering in the Cu-Mn system has been estimated by Yosida to be 2.1×10^{-12} erg for a manganese ion of spin $\frac{5}{2}$. This value of J is higher but of the same order of magnitude as the value of 1.5×10^{-12} found for the zinc-manganese alloys reported in the present paper. An estimate for J from the suppression of the superconducting critical temperature for the Al-Mn system[3] yields a value of 0.23×10^{-12} erg for an estimated spin value for manganese of $\frac{1}{2}$. The fact that manganese possesses a very small moment, if any, when dissolved in aluminum[15] would indicate that the mechanism responsible for the suppression of T_c in the Al-Mn system is due to a cause other than that responsible for the observed effects in the Zn-Mn system.

Calculations of the exchange integral from the magnetic properties of dilute zinc alloys containing small concentrations of manganese have been made from the Néel temperature, the magnetoresistivity as a function of field strength and the suppression of the superconducting critical temperature. The theoretical equations used for these exchange integral determinations are all based on a conduction electron-paramagnetic ion exchange interaction of the form $2J$ S.s. The analysis of the experimental results based on this interaction all yield the same value for the exchange integral which

[15] E. W. Collings and F. T. Hedgcock, Phys. Rev. 126, 1654 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

could lead one to conclude that the nature of the magnetic interaction that leads to these phenomena is of the same physical origin. Graphical analysis of the low-temperature resistance characteristics of dilute Zn-Mn alloys, assuming the sum of two terms; one due to magnetic ordering and one due to resonance scattering results in a value of J in agreement with that derived from other transport phenomena. This then suggests a method of separating the resistance minimum from total resistivity in the presence of a resistance maximum produced by magnetic ordering.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge discussions with Dr. M. Merriam and Dr. W. B. Muir and correspondence with Dr. I. Rizzuto, Dr. T. S. Dugdale, and Dr. M. Bailyn.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Covalency in Crystal Field Theory: $KNiF_3$

R. E. WATSON

Bell Telephone Laboratories, Murray Hill, New Jersey

AND

A. J. FREEMAN

National Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts*

(Received 13 November 1963; revised manuscript received 13 December 1963)

The theory of covalency in crystal field phenomena is examined, using, as example, the Ni-F_6 complex in $KNiF_3$. The Hund-Mulliken-Van Vleck molecular orbital-linear combination of atomic orbitals treatment is followed. The role of the antibonding *and* bonding electrons in the complex is discussed from a multielectron point of view. The exact self-consistent one-electron Hamiltonian is discussed in some detail. Emphasis is placed on elucidating the source and nature of the covalent effects appropriate to the various physical phenomena. We find that it is the covalent mixing of those *bonding* electrons having no antibonding partners which contribute to all experimental observables (including the crystal field splitting 10 Dq, transferred hyperfine interactions, neutron magnetic form factors, and superexchange interactions). This view of covalency differs markedly from the one followed by Sugano and Shulman, in that the covalency of the *antibonding* electrons, which are assigned the sole role in their approach, is totally irrelevant. Quantitative numerical estimates (using approximations to the exact Hamiltonian) are given for the two models of the covalent effects in $KNiF_3$, i.e., "unpaired" bonding and antibonding; they are shown to differ strongly. The relative roles of overlap and covalency are discussed; covalency is found to play an important but by no means dominant role. Numerical agreement between the present inexact cluster theory and experiment is found to be poor. The various sources of this disagreement are reviewed.

I. INTRODUCTION

CRYSTAL field theory[1,2] has had a long and varied history: when treated as a semiempirical theory, with the crystal field splitting 10 Dq considered as an adjustable parameter, it has been highly successful in fitting experimental data; when considered as a fundamental theory for the behavior of transition metal ions in crystalline fields, it has been strikingly unsuccessful in predicting, from first principles, the fundamental parameter 10 Dq. Following the pioneering computations of Van Vleck[3] and Polder,[4] a series of theoretical investigations[5–9] succeeded in pinpointing the basic shortcomings of the theory, and indicated the need for a multielectron many-center molecular approach.

The recent work of Sugano and Shulman[10] (henceforth denoted as S&S III), representing the most detailed computations undertaken to date, attempted to obtain a quantitative basis for the theory by including all the terms in the ionic model considered by their predecessors, as well as the effect of metal ion-ligand covalent mixing. A cluster model consisting of the metal ion and its nearest ligand neighbors in an external Madelung-like potential was invoked. Considering $KNiF_3$, they obtained a theoretical cubic field splitting parameter (10 Dq) for Ni which gave the first quantitative indication that covalency plays an important role in the crystal field interactions of salts as highly ionic as

* Supported by the U. S. Air Force Office of Scientific Research.

[1] H. Bethe, Ann. Physik 3, 133 (1929).
[2] J. H. Van Vleck, Phys. Rev. 41, 208 (1932).
[3] J. H. Van Vleck, J. Chem. Phys. 7, 72 (1939).
[4] D. Polder, Physica 9, 709 (1942).
[5] W. H. Kleiner, J. Chem. Phys. 20, 1784 (1952).
[6] H. S. Jarrett, J. Chem. Phys. 31, 1579 (1959).
[7] Y. Tanabe and S. Sugano, J. Phys. Soc. Japan 11, 864 (1956).
[8] J. C. Phillips, Phys. Chem. Solids 11, 26 (1959).
[9] A. J. Freeman and R. E. Watson, Phys. Rev. 120, 1254 (1960). This reference also contains a useful review of the theoretical investigations prior to the work of Ref. 10.
[10] S. Sugano and R. G. Shulman, Phys. Rev. 130, 517 (1963); henceforth denoted as S&S III.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$KNiF_3$. In addition, they considered the role of the computed covalent mixing in the F^{19} transferred hyperfine interactions.[11]

While the role of covalency in superexchange,[12] and transferred hyperfine effects[11,13–17] have been particular objects of study, overlap and covalent mixing significantly affect almost every observable in an iron series salt. Covalent behavior must be understood in detail before matters such as the theory of superexchange can be put on any sort of quantitative basis.

Sugano and Shulman's modification of crystal field theory to include covalency in the cluster model may be viewed as an extension of the traditional approach to crystal field effects in ionic salts. In $KNiF_3$, the familiar triply (t_{2g}) and doubly (e_g) degenerate $3d$ orbitals, describing the Ni^{2+} ion in its ground-state configuration ($t_\uparrow^3 t_\downarrow^3 e_\uparrow^2$), in the ionic model, become antibonding orbitals in covalent theory. These orbitals are shown schematically in Fig. 1. Also included are the bonding orbitals, which are formed predominantly from ligand $2s$ and $2p$ functions. Since their energy is lower than that of the antibonding orbitals, they are completely occupied. According to S&S, it is the covalency of the antibonding electrons which must be considered in order for the experimental observables to be reproduced. Thus, in the theory of the transferred hyperfine effects only the covalency of (spin) unpaired antibonding $3d$ electrons of majority spin contribute, while the crystal field splitting is associated with the $3d$ antibonding electrons of minority spin. Sugano and Shulman utilized the conventional approach of treating electrons of either spin identically and therefore obtained a common estimate of this covalent mixing for the two phenomena. In this approach, the bonding electrons are considered to play no role (other than to provide orthogonal partners for the antibonding electrons).

In this paper, the theory of covalency in crystal field phenomena is examined using the cluster model for $KNiF_3$ as example. We find the covalent mixing of those *bonding* electrons which have no antibonding partners, to be the appropriate manifestation of covalency in experiment. Unlike the antibonding version described above, it is the covalent mixing of the same (unpaired bonding) electrons (of minority spin in $KNiF_3$) which contributes to both $10\,Dq$ and the transferred hyperfine interactions and, in fact, to all experimental observables. The covalency (but not the overlap) of the antibonding electrons is entirely irrelevant because it is exactly compensated by the covalency of their bonding partners. This agrees with several features of the role of bonding electrons in transferred hyperfine theory which have been described in the context of the Heitler-London method by Keffer *et al.*[13] and for the molecular orbital (MO) method by Clogston *et al.*[15] These matters are important for if we solve the Hartree-Fock equations for the two types of covalent mixing we find them to differ markedly. This difference does not imply different orbitals for different spin in the sense of unrestricted Hartree-Fock (UHF) theory, because the antibonding covalency (which has *no* physical or variational meaning) may have any value—including that appropriate to the unpaired bonding electrons. (We briefly discuss, for the case of Cr^{3+}, the implications of "UHF covalent mixing.")

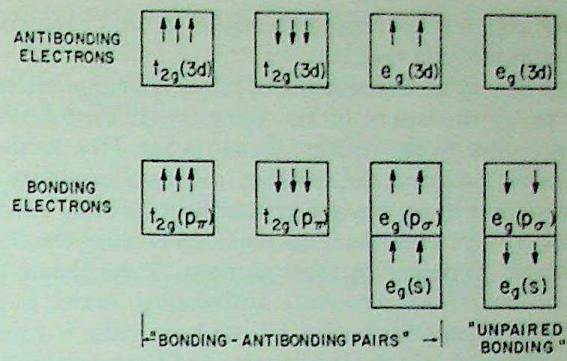


FIG. 1. A schematic representation of the antibonding and bonding electrons grouped into bonding-antibonding pairs and unpaired bonding electrons. For each, the predominant atomic orbital, from which the molecular orbital is formed, is indicated in parentheses; the F^- ligand $2s$, $2p_\sigma$, and $2p_\pi$ orbitals are denoted simply by s, p_σ, and p_π.

Two basic aspects of this crystal field theory are considered in what follows: (1) the quantitative implications of the approximations made in the treatment of a Hamiltonian, its matrix elements and related matters; (2) the source and nature of the covalent and overlap effects appropriate to the various physical phenomena. We concentrate on the latter in some detail in the present investigation and leave some aspects of the quantitative calculations to a future paper. Our studies of covalency are described in terms of the occupied one-electron orbitals of the cluster. An alternate approach to crystal field problems consists of dealing with the unoccupied antibonding orbitals, i.e., the antibonding holes (which are associated with the unpaired bonding electrons). Such a treatment of covalent mixing, and the resulting estimates of $10\,Dq$ and other experimental parameters, is exactly equivalent to that associated with the unpaired bonding orbital treatment provided that certain restrictions (to be discussed later) are maintained. For this reason, we will consider only one approach in detail but will make

[11] K. Knox, R. G. Shulman, and S. Sugano, Phys. Rev. **130**, 512 (1963); denoted as K, S&S II. R. G. Shulman and S. Sugano, *ibid.* **130**, 506 (1963); denoted as S&S I. R. G. Shulman and K. Knox, *ibid.* **119**, 94 (1960) [in particular, see the discussion between Eqs. (1) and (4)].

[12] See, e.g., P. W. Anderson, Solid State Phys. **14**, 99 (1963); of the references cited therein, see, in particular, J. Kondo, Progr. Theoret. Phys. (Kyoto) **18**, 541 (1957).

[13] F. Keffer, T. Oguchi, W. O'Sullivan, and J. Yamashita, Phys. Rev. **115**, 1553 (1959).

[14] M. Tinkham, Proc. Roy. Soc. (London) **A236**, 549 (1956).

[15] A. M. Clogston, T. P. Gordon, V. Jaccarino, M. Peter, and L. R. Walker, Phys. Rev. **117**, 1222 (1960).

[16] A. J. Freeman and R. E. Watson, Phys. Rev. Letters **6**, 343 (1961).

[17] W. Marshall and R. Stuart, Phys. Rev. **123**, 2048 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

brief contact with the unoccupied orbital prediction when describing results.

In Sec. II we describe the covalent mixing of interest to us. A discussion of the one-electron Hamiltonian appropriate to the system follows (in Sec. III). Sections IV and V discuss the covalent mixing appropriate to transferred hyperfine and crystal field splitting effects, respectively. The role of the unpaired bonding electrons becomes apparent in these sections. Throughout, the traditional (and incorrect) antibonding model is examined in parallel. In Sec. VI, we have cause to return to the matter of our Hamiltonian and the crucial role it plays in causing the two types of computed covalent mixing to differ. Since we use an approximate Hamiltonian, as in Ref. 10, certain "self-energy" problems arise which lead to serious internal inconsistencies in the results. The nature of these inconsistencies are explored. Their resolution requires the use of a better approximation to the true one-electron Hamiltonian, a matter reserved for the future. The various matrix elements appearing in our calculations are discussed, and their numerical values are reported, in Sec. VIII. Results, a discussion of their implications for other physical parameters, and conclusions follow.

II. THE MOLECULAR ORBITAL APPROACH

The molecular orbital approach is by now well known; hence, we will but briefly recapitulate some of the definitions and ideas required by us here. The antibonding molecular orbitals are defined by

$$\begin{aligned}\Psi_e{}^A &\equiv N_e(\varphi_e-\lambda_s\chi_{2s}-\lambda_\sigma\chi_{2p\sigma}),\\ \Psi_t{}^A &= N_t(\varphi_t-\lambda_\pi\chi_{2p\pi}),\end{aligned} \tag{1}$$

where the N_i are normalization constants, the φ's are $Ni^{2+}3d$ orbitals of t (xy, yz, and zx) and e (x^2-y^2 and $3z^2-r^2$) symmetry, and the χ's are appropriate linear combinations[11] of $2s$, $2p\sigma$, or $2p\pi$ atomic orbitals (ψ_i's) associated with the six nearest-neighbor F^- ions. The λ_i's are the covalent mixing parameters. The normalization constants are given by

$$\begin{aligned}N_e &= [1-2\lambda_s S_s-2\lambda_\sigma S_\sigma+\lambda_s{}^2+\lambda_\sigma{}^2+2\lambda_s\lambda_\sigma S_{s\sigma}]^{-1/2};\\ N_t &= [1-2\lambda_\pi S_\pi+\lambda_\pi{}^2]^{-1/2},\end{aligned} \tag{2}$$

where an S_i is the overlap integral between χ_i and the φ_{3d} orbital with which it is being combined, and $S_{s\sigma}$ is the overlap integral between χ_s and χ_σ; it need not be zero valued, as is often assumed.

One can alternatively speak of the covalent mixing associated with the bonding molecular orbitals,

$$\begin{aligned}\Psi_{es}{}^B &= N_{es}'(\chi_{2s}+\gamma_s\varphi_e+\gamma_{s\sigma}\chi_{2p\sigma}),\\ \Psi_{e\sigma}{}^B &= N_{e\sigma}'(\chi_{2p\sigma}+\gamma_\sigma\varphi_e+\gamma_{\sigma s}\chi_{2s}),\\ \Psi_t{}^B &= N_t'(\chi_{2p\pi}+\gamma_\pi\varphi_t),\end{aligned} \tag{3}$$

where

$$\begin{aligned}N_{es}' &= [1+2\gamma_s S_s+2\lambda_{s\sigma}S_{s\sigma}+2\gamma_s\gamma_{s\sigma}S_\sigma+\gamma_s{}^2+\gamma_{s\sigma}{}^2]^{-1/2},\\ N_{e\sigma}' &= [1+2\gamma_\sigma S_\sigma+2\lambda_{\sigma s}S_{s\sigma}+2\gamma_\sigma\gamma_{\sigma s}S_s\\ &\qquad +\gamma_\sigma{}^2+\gamma_{\sigma s}{}^2]^{-1/2},\\ N_t' &= [1+2\gamma_\pi S_\pi+\gamma_\pi{}^2]^{-1/2}.\end{aligned} \tag{4}$$

The bonding and antibonding orbitals are to be orthonormal, so to lowest order

$$\begin{aligned}\lambda_s &= \gamma_s+S_s,\\ \lambda_\sigma &= \gamma_\sigma+S_\sigma,\\ \lambda_\pi &= \gamma_\pi+S_\pi,\\ \gamma_{s\sigma} &= -\gamma_{\sigma s}-S_{\sigma s}.\end{aligned} \tag{5}$$

The bonding and antibonding orbitals are to be eigenfunctions of the Hartree-Fock (H-F) equations for a ($Ni-F_6$) cluster in $KNiF_3$, i.e., they are to satisfy

$$h\Psi_i = \epsilon_i\Psi_i, \tag{6}$$

where h is the one-electron self-consistent H-F Hamiltonian. In the present paper, we will follow Sugano and Shulman and approximate the effects of the crystal external to the cluster by an electrostatic Madelung-like potential V_{ext}, i.e., we assume that there are no covalent or overlap effects, between the cluster and its environment, which significantly affect the metal-ligand covalency of interest here. This approximation is necessary in order to make the problem tractable. (One can improve on the approximation by making V_{ext} the most accurate description of environmental effects obtainable with a local *effective* potential, but it is well to note that there are shortcomings in such an approach—shortcomings which are particularly significant when one considers superexchange or any other interaction between ions in different clusters.) The use of such a V_{ext} is compatible with a self-consistent H-F treatment for the cluster itself.

One automatic result of a self-consistent H-F treatment is that a set of orthonormal one-electron eigenfunctions is obtained. We are thus supplied with a test of the seriousness of any approximation we make in our Hamiltonian (or in evaluating matrix elements) which causes us to fail to have a proper self-consistent theory. The test consists of independently obtaining an antibonding orbital, Ψ^A, and its bonding partner, Ψ^B, and checking their orthogonality. We will see that the present approximate cluster theory does not always meet this orthogonality requirement.

We will solve Eq. (6) for Ψ within a very limited function space, namely, one spanned by the molecular orbitals constructed from Ni^{2+} and ligand-free ion orbitals. (One well might wish to allow greater variational freedom in a future calculation.) However, what is important to us here, is that the limited nature of this space in no way effects eigenfunction orthogonality or the self-energy problem to be discussed in Sec. VI. Let us

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

now consider, in detail, the nature of the one-electron Hamiltonian.

III. THE ONE-ELECTRON HAMILTONIAN

We are interested in obtaining the exact one-electron H-F Hamiltonian, appropriate to Eq. (6), derived from a many-electron Hamiltonian consisting of kinetic and electrostatic terms, where the environment external to the $(Ni-F_6)$ cluster is approximated by a Madelung-like one-electron electrostatic potential $V_{ext}(r_i)$. The process is made trivial by the fact that the two H-F many-electron states of interest to us are single determinants constructed from orthonormal one-electron orbitals (Ψ^A's and Ψ^B's). For such states the one-electron Hamiltonian h_i, for electron orbital i, is simply

$$h_i(r_1)=\tfrac{1}{2}\nabla_1^2+\sum_\alpha\left(-\frac{Z_\alpha}{r_{\alpha 1}}\right)+V_{ext}(r_1)+\sum\int\Psi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\Psi_j(r_2)d\tau_2 \quad (7)$$

in atomic units. The α summation (over nuclear potential contributions) is limited to the one Ni and six F nuclei for we have already introduced the potential accounting for all electrons and nuclei (presumed) external to the cluster. The P_{12} operator permutes coordinates r_1 and r_2 giving us the interelectronic exchange terms, which occur only between electrons of common spin. The sum over interelectronic terms is limited to the cluster and consists of contributions from all electrons in the cluster including the ith electron itself. This is allowed in a self-consistent H-F theory because the ith electron Coulomb contribution is exactly canceled by its exchange term. We will have occasion to return to the matter of including or excluding an electron's self-energy contribution to Eq. (7) later.

We now with to re-express the interelectronic potential of h in terms of one-electron functions localized on the seven nuclei in the cluster, for it is in terms of these that the potential, of Eq. (7), can be evaluated. In this simple procedure the Ψ's are expressed as linear combinations of the local atomic orbitals (LCAO's), i.e., Ψ is broken up into Ni φ and ligand χ_i orbitals [with the χ_i's themselves expressed in terms of local orbitals (ψ_i's) at the various fluorine sites], whereupon Eq. (7) may be written as

$$h_i(r_1)=-\tfrac{1}{2}\nabla_1^2+V_{ext}(r_1)+V_m(r_1)+V_L(r_1)+V_S^{(1)}(r_1)+V_\gamma^{(1)}(r_1)+V_S^{(2)}(r_1)+V_\gamma^{(2)}(r_1)+V_{SL}(r_1). \quad (8)$$

As we shall see, the terms of the second line account for the important fact that the local one-electron functions of one center are nonorthogonal with those of another (i.e., $S\neq 0$) and that there is covalent mixing.

Let us now consider the individual terms separately. V_m: This term is simply the H-F nuclear and interelectronic Coulomb plus exchange potential for an isolated Ni^{2+} ion;

$$V_m(r_1)=-\frac{Z_{Ni}}{r_1}+\sum_{\substack{j=\text{all Ni}^{2+}\\ \text{electrons}}}\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\varphi_j(r_2)d\tau_2. \quad (9)$$

V_L: These are similar contributions from the ligand ions;

$$V_L(r_1)=\sum_{\alpha=1}^{6}\left\{\frac{-Z_\alpha}{r_{1\alpha}}+\sum_{\substack{j=\text{all electrons}\\ \text{on ion }\alpha}}\int\psi_{j,\alpha}^*(r_2)\frac{1-P_{12}}{r_{12}}\times\psi_{j,\alpha}(r_2)d\tau_2\right\}. \quad (10)$$

The above terms would make up the interelectronic contributions coming from the cluster if there were neither covalent mixing *nor* any overlap between orbitals on different sites. The omission of overlap terms may be thought of as yielding a potential of the form one would have when the ions of the cluster are infinitely separated. It is important to note that such a potential [actually the full first line of Eq. (8)] was used by Sugano and Shulman in their work. The consequences of this approximation will be discussed later.

Next we consider the overlap contributions to the potential. $V_S^{(1)}(r_1)$ and $V_S^{(2)}(r_1)$: The most obvious overlap effects are associated with metal-ligand nonorthogonality. In the case of zero covalency, this nonorthogonality is resolved by having $\lambda_i=S_i$ while $\gamma_i=0$ for the bonding-antibonding pairs. Setting λ_i and γ_i to these values, we can inspect the overlap terms arising in Eq. (8). There are terms linear in S, i.e.,

$$V_S^{(1)}(r_1)=-\sum_j 2S_j\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2, \quad (11)$$

where the sum is over all bonding-antibonding pairs. Note that Eq. (11) is written in terms of the multicenter ligand molecular orbitals, χ_j. There are also higher order overlap terms in S_j, i.e.,

$$V_S^{(2)}(r_1)=+\sum_j S_j^2\left\{\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\varphi_j(r_2)d\tau_2+\int\chi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2\right\}+O(S^3). \quad (12)$$

These second-order terms are of the same order of importance as $V_S^{(1)}$ which involves an S_j times an overlap charge density. Note that a φ_s appears twice in the summations in these equations to account for its two partners, χ_s and $\chi_{s'}$. Third and higher order terms are, of course, less important.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

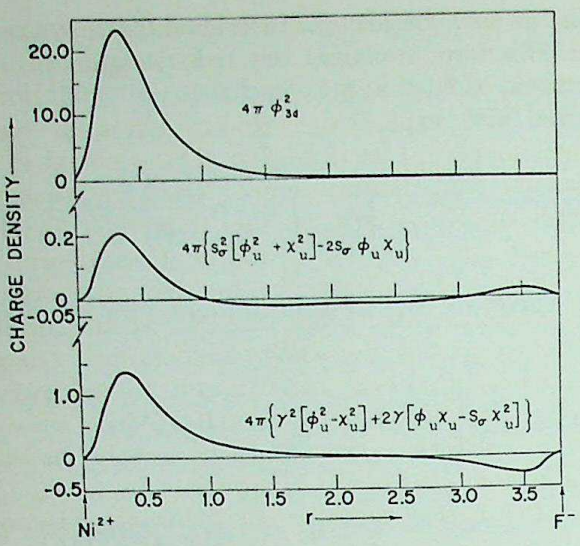


FIG. 2. The charge density of $4\pi\varphi_u^2$ of the $u(=3z^2-r^2)$ orbital compared with the p electron overlap charge density $4\pi\{S_\sigma^2[\varphi_u^2+\chi_\sigma^2]-2S_\sigma\varphi_u\chi_\sigma\}$ and covalent charge density $4\pi\{\gamma^2[\varphi_u^2-\chi_\sigma^2]+2[\varphi_u\chi_\sigma-S_\sigma\chi_\sigma^2]\}$ plotted along the z axis from Ni^{2+} to F^-. Densities are all in the same units, normalized such that $\int\varphi_{3d}^2(\mathbf{r})d\tau=1$; the free ion S value was used (see Table VII) and $S+\gamma$ was chosen so as to reproduce the experimental (Ref. 11) anisotropic F^- hyperfine term.

In Fig. 2 we plot the overlap charge density, giving rise to the sum of the $V_S^{(1)}$ and $V_S^{(2)}$ terms, along a metal-ligand axis, and for comparison, the charge density associated with an antibonding electron prior to the introduction of overlap and covalency [i.e., just $|\varphi_4(\mathbf{r})|^2$]. We see that the antibonding charge density has been delocalized, by the addition of the overlap terms, in the sense that there has been a buildup of charge on the ligand (which is one of six so affected), but, there has also been an increase of charge on the metal as well. These increases have been compensated for by a decrease of charge due to the $-2S_i[\varphi_i^*(\mathbf{r})\chi_i(\mathbf{r})]$ term, in the region between the ions. In other words, one is not dealing with a shift in charge (or spin) simply from one ion to another but onto the ions from the region between. The often cited picture of overlap (and covalent) effects delocalizing charge and spin off the metal ion is, therefore, not valid, for we are dealing with a process which is more complicated. Whether the overlap effect *acts as if* there is a shift onto, or off, the metal ion, depends strongly on the nature of the operator whose expectation value is in question.

V_{SL}: These are additional overlap potential contributions due to the nonorthogonality between ligand orbitals. First, ligand-ligand ($\psi-\psi$) overlap enters into the evaluation of the second line of Eq. (12) (where we have χ^2), and elsewhere. We will include the latter terms as part of V_{SL} but will not write them out explicitly. It should be noted that the normalization of the χ_i's as defined in S&S I [Eq. (2.4)], is affected by this nonorthogonality; however, for the purposes of this paper, we assume their normalization covention. Secondly, $\chi_s-\chi_\sigma$ nonorthogonality leads to additional potential terms, the lowest order ones being of the form

$$S_{s\sigma}{}^2\left[\int\chi_\sigma^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_\sigma(r_2)d\tau_2+\int\chi_s^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_s(r_2)\right]$$
$$-2S_{s\sigma}\int\chi_\sigma^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_s(r_2)d\tau_2 \quad (13)$$

similar to what we have seen in Eqs. (11) and (12). Finally, the simultaneous nonorthogonality of φ_e with χ_s and χ_σ, which leads to terms involving the $\chi_\sigma\chi_s$ charge density, have been arbitrarily included in V_{SL}. The lowest order term is of the form

$$2S_sS_\sigma\int\chi_s^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_\sigma(r_2)d\tau_2 \quad (14)$$

and is of higher order than the contributions written out explicitly in Eqs. (11), (12), and (13), i.e., it is the product of two overlap integrals times an overlap charge density potential. Finally we come to covalent contributions. $V_\gamma^{(1)}(r_1)$, $V_\gamma^{(2)}(r_2)$: These are most simply obtained by evaluating the interelectronic term Eq. (8) for nonzero γ_j (with $\lambda_j=\gamma_j+S_j$), and subtracting off the overlap terms already accounted for in $V_S^{(1)}$ and $V_S^{(2)}$. One obtains linear terms of the form

$$V_\gamma^{(1)}(r_1)=-\sum_{\substack{j=\text{antibonding}\\ \text{orbitals}}}2\gamma_j\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2$$
$$+\sum_{\substack{j=\text{bonding}\\ \text{orbitals}}}2\gamma_j\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2. \quad (15)$$

It is immediately apparent that contributions from members of an occupied bonding-antibonding pair cancel one another, leaving us *only* with contributions coming from *unpaired bonding orbitals*, i.e.,

$$V_\gamma^{(1)}(r_1)=+\sum_{\substack{j=\text{unpaired bonding}\\ \text{orbitals}}}2\gamma_j\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\times\chi_j(r_2)d\tau_2. \quad (16)$$

Such cancellations also occur for the covalent mixing involving (a nonzero) $\gamma_{\sigma s}$ in a $\Psi_s{}^B-\Psi_\sigma{}^B$ pair. A similar cancellation occurs in second order giving us

$$V_\gamma^{(2)}(r_1)=+\sum_{\substack{j=\text{unpaired}\\ \text{bonding}}}\left\{\gamma_j{}^2\left[\int\varphi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\varphi_j(r_2)d\tau\right.\right.$$
$$\left.-\int\chi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2\right]$$
$$\left.-2\gamma_jS_j\int\chi_j^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_j(r_2)d\tau_2\right\}$$
$$+\text{higher order terms.} \quad (17)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The cancellation illustrated in detail above follows simply from a well-known property of determinantal functions: Any linear combination of occupied orbitals in a Slater determinant leaves unchanged the expectation value of an operator—and hence does not affect the prediction of any physical observable.

The first line of Eq. (17) is associated with an obvious physical effect, namely, the transfer of charge from ligand to metal ions due to covalent mixing. The number of electrons so transferred is simply

$$N = \sum_{i=\text{unpaired bonding orbitals}} \gamma_i^2. \tag{18}$$

Antibonding-bonding pairs do not contribute to this transfer because the covalency of one is compensated for by that of the other, as is required, to second order, by the orthogonality conditions [Eqs. (5)]. The third line of Eq. (17) is coupled with $V_\gamma^{(1)}$ which is $2\gamma_j$ times an overlap potential. The two terms represent a first-order (in γ_j) shift of charge off the ligands into the space between ligands and the Ni^{2+} ion. It should be noted that as the unpaired bonding orbitals are of minority spin, the covalent terms involve a shift of minority spin onto the metal ion, effectively reducing the net spin localized at the Ni^{2+} site. These covalent charge density terms are plotted for a bonding $2p\sigma$ electron in Fig. 2 (their negative should be taken when considering spin effects) where γ was chosen so as to reproduce the experimental[11] anisotropic transferred hyperfine interaction. This choice causes these covalent effects to be greater than the overlap terms for the same e_g electrons. Such a dominance need not always occur. Note that there are also t_{2g} and Ni closed shell overlap charge, but *not* spin, density effects. The overlap and covalent e_g terms act cooperatively in their contributions to the Ni^{2+} charge and F^- spin densities while making opposing contributions to the Ni^{2+} spin and F^- charge behaviors.

We have seen a covalent shift of charge associated with unpaired bonding electrons and no such shift for bonding-antibonding pairs. This fact is of great interest to us for reasons beyond the mere matter of estimating a potential, for it indicates the profoundly different nature of the two types of covalent mixing. The mixing of a bonding-antibonding pair conserves the charge on the metal ion while that of an unpaired bonding electron involves the shift of charge from ligands to metal. This observation suggests that the computed covalent mixing, i.e., the γ_i, differs for the two cases and is the reason for our preoccupation with bonding electron behavior in this paper.

In the above discussion of the overlap and covalent contributions to the Hamiltonian, h, we deliberately neglected to specify exactly which orbitals were so involved. As covalent terms only occur in low order for unpaired bonding orbitals we need only consider those electrons which can enter into bonding with the open $3d$ shell; i.e., the $2s$, $2p$, and $1s$ shells of the ligands. The $2s$ and $2p$ electrons are expected to make small contributions, and the $1s$ electrons insignificant contributions, to $V_\gamma^{(1)}$ and $V_\gamma^{(2)}$.

We also expect the overlap terms to make a more significant contribution to the potential terms than do the covalent terms. Now, any shell which significantly overlaps another ion shell should be included in $V_S^{(1)}$, $V_S^{(2)}$, and/or V_{SL}. Hence, when including overlap terms in the potential one must inspect terms other than the Ni $3d$ and ligand $2s$ and $2p$ electrons, to which consideration is normally limited. In particular, one should, at the minimum, explore the role of the Ni $3s$ and $3p$ shells since their overlap with the ligand electrons is only moderately smaller than that of the $3d$ orbitals.

In the present paper we will follow the Sugano and Shulman treatment and mostly limit ourselves to a Hamiltonian (denoted by h_0) consisting of the first line of Eq. (8) alone for two reasons. First, we believe it imperative to fully explore the implications of their approximations, approximations which at first glance seem to yield excellent agreement with experiment, before going on to a more complete (and more complicated) treatment. Secondly, a more complete treatment appears to require far greater care in the numerical evaluation of matrix elements than does a treatment based on h_0 alone. Greater care is not only necessary in the handling of overlap and covalent terms but also in the evaluation of the h_0 matrix elements. Approximations which were apparently adequate for the simple case must be re-examined before one attempts the more detailed treatment; this is due, in part, to the severe numerical differencing which occurs in the course of evaluating matrix elements. Our reasons for being cautious on this matter will become obvious when we inspect the results of the present paper.

As Sugano and Shulman have emphasized, the use of h_0 can be viewed as the first iteration of a self-consistent theory for the $(Ni-F_6)$ covalent mixing problem. This is so, providing that overlap as well as covalent terms are included in a later iteration. As we shall see, the results of Sec. VIII suggest that a proper evaluation of covalent effects cannot be had without the inclusion of the overlap and covalent terms in the Hamiltonian [i.e., the second line of Eq. (8)]. It is our intention to investigate the effect of such terms in a future paper.

In going beyond the use of h_0, one is involved in an ever increasing commitment to details of the cluster model, details which do not reflect the situation in the actual crystal. The use of h_0 and the MO's as defined earlier for the cluster already imply a strong commitment.

IV. UNPAIRED SPIN COVALENCY AND TRANSFERRED HYPERFINE EFFECTS

We now consider the role of covalent mixing in transferred hyperfine effects. In doing this, we assume that the orthogonality relations of Eq. (5) hold and

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

that orbitals differing in spin may have different covalent parameters.

The isotropic F^{19} hyperfine parameter is considered to be primarily associated[14,11] with the unpaired $2s$ spin induced by nonorthogonality and covalency on the F^- site. After a variety of other contributions[16,17,11] are accounted for, the parameter is written as the product of the hyperfine constant associated with a single F^- $2s$ orbital times the fraction, f_s, of induced unpaired $2s$ spin character at the F^- site. Making the standard assumption of constant covalency for electrons of either spin, the value of f_s in $KNiF_3$ is

$$f_s=\lambda_s^2 N_s^2/3 \tag{19}$$

and arises from the antibonding electrons alone. The $\frac{1}{3}$ accounts for the fact that λ is defined in terms of molecular orbitals and that we are interested in the $2s$ character induced on a single F^- site (via both antibonding $e_\uparrow$ orbitals). The anisotropic hyperfine interaction is complicated by the presence of (1) both p_σ and p_π conbributions and (2) a substantial correction for classical spin dipolar effects (which must be made prior to estimating covalency from experiment). The p-electron interaction is entirely spin dipolar (i.e., not orbital) and given the experimental covalent hyperfine parameter, the fractional difference $(f_\sigma-f_\pi)$between induced p_σ and p_π behavior can be inferred.

The coventional assumption of common covalency for either spin leads to $f_\pi=0$ for $KNiF_3$ since all t orbitals are occupied in the ground state. Experiment, therefore, would seem to yield information concerning the $2s$ and $2p_\sigma$ covalent mixing in the antibonding $e_\uparrow$ orbitals from the observed isotropic and antisotropic interactions.

The above discussion assumes that the transferred hyperfine interaction is entirely due to the spin density associated with a ligand's ψ orbitals. However, the tails of the free ion Ni φ_{3d} (and neighboring ligand) orbitals overlap into the region of a F nucleus and contribute[16,17] to the spin density and hyperfine interaction via the φ^2 and $\varphi\chi$ terms. These terms are not negligible.[16] Unfortunately the use of free ion orbitals is inappropriate for this evaluation since these orbitals do not take account of the presence of neighboring ion nuclei. We will encounter similar problems, introduced by inappropriate orbital tail behavior, when discussing orbital reduction effects in Sec. IX. These matters will be ignored in the remainder of this section.

Let us now consider how the picture changes when one accounts for all covalent mixing while maintaining the necessary orthogonality of bonding and antibonding orbitals to second order in S_i, χ_i, and γ_i. We will again consider the $2s$ effects as they are the least complicated. The implications for anisotropic effects will then be immediately apparent.

Consider the contributions to f_s from the bonding $e_\uparrow$ and $e_\downarrow$, and antibonding $e_\uparrow$ electrons. Writing out the normalization constant in Eq. (17) and making use of Eq. (5), the antibonding contribution becomes

$$f_{s\uparrow}{}^A=\tfrac{1}{3}\{S_s{}^2+2S_s\gamma_{s\uparrow}+\gamma_{s\uparrow}{}^2\}+0(S^3). \tag{20}$$

(We will henceforth use superscripts A and B to denote antibonding and bonding, respectively, whenever appropriate for clarity.) In the absence of covalent bonding and $\chi_s-\chi_\sigma$ overlap, the f_s contributions are $+1$ for spin $\uparrow$ and -1 for spin $\downarrow$, and hence cancel. Covalency affects these f_s contributions via (1) the normalization of the $\Psi_{es}{}^B$ orbitals [cf. Eq. (4)] and (2) the χ_{2s} term of $\Psi_{e\sigma}{}^B$. With differing covalency, these contributions need no longer cancel. The individual spin contributions are

$$f_{s\uparrow}{}^B=\tfrac{1}{3}N_{s\uparrow}{}^2+\tfrac{1}{3}N_{\sigma\uparrow}{}^2\gamma_{\sigma s\uparrow}{}^2$$
$$=\tfrac{1}{3}\{1-2\gamma_{s\uparrow}S_s-\gamma_{s\uparrow}{}^2+S_{s\sigma}{}^2\}+0(S^3)+\cdots, \tag{21}$$
$$f_{s\downarrow}{}^B=-\tfrac{1}{3}\{1-2\gamma_{s\downarrow}S_s-\gamma_{s\downarrow}{}^2+S_{s\sigma}{}^2\}+0(S^3)+\cdots. \tag{22}$$

Combining all these contributions results in

$$f_s=\tfrac{1}{3}\{S_s{}^2+2\gamma_{s\downarrow}S_s+\gamma_{s\downarrow}{}^2\}+0(S^3). \tag{23}$$

Similarly, one obtains

$$f_\sigma=\tfrac{1}{3}\{S_\sigma{}^2+2S_\sigma\gamma_{\sigma\downarrow}+\gamma_{\sigma\downarrow}{}^2\}+0(S^3). \tag{24}$$

In other words, if bonding and antibonding orbitals of one spin are simultaneously occupied, there will be an $S_i{}^2$ overlap contribution, but their covalent terms cancel exactly. Hence, it is the covalency of the *bonding* $e_{s\downarrow}{}^B$ electrons which contribute to the isotropic transferred hyperfine interaction in $KNiF_3$, and if there is any question as to the constancy of $2s$ covalent mixing within the set of e orbitals, it is the mixing estimated for the $e_\downarrow{}^B$ electrons which must be compared with experiment.

The coventional cancellation associated with spin pairing has been lost, but the more fundamental pairing cancellation between bonding and antibonding orbitals has been seen. The immediate effects of this for the anisotropic F^- hyperfine interaction are (1) that $t_\uparrow$ and $t_\downarrow$ electrons cannot contribute, no matter what their respective covalencies, since antibonding-bonding pairs of both spins are occupied; (2) that there will be S^2, but no γ contributions from the majority spin $e_\uparrow{}^B$ and $e_\uparrow{}^A$ electrons; and (3) that covalent contributions come solely from unpaired $e_\downarrow{}^B$ electrons of minority spin. It is clear from the identical form of Eqs. (20) and (23) that the usual phenomenological approach, which uses the antibonding expression [Eq. (20)], actually determines[15] the unpaired bonding covalent mixing parameter.

Complications occur for almost empty d shell ions such as $Cr^{3+}(t_\uparrow{}^A)^3$ which has but three antibonding d-like electrons. There are no e^A electrons in the ground state and differences in $e_\uparrow{}^B$ and $e_\downarrow{}^B$ covalent mixing can lead, for example, to an isotropic F^- hyperfine term which would be zero in the traditional analysis.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Transferred hyperfine effects have been observed[18] in K_2NaCrF_6; the isotropic term is almost zero valued. This either implies that the e^B covalency differences are very small here or that an accidental near cancellation of a number of contributions has occurred. Neutron diffraction experiments on such a system could test for e electron character in the cluster's spin density.

It must be stressed that any computed differences in $e_\uparrow{}^B - e_\downarrow{}^B$ covalency, for this case, come from a different source than do the differences between the covalency of an occupied Ni^{2+} bonding-antibonding pair of one spin and that of an unpaired bonding orbital of the other. The covalent "unpairing" of the Cr^{3+} orbitals has its source in the unbalanced exchange terms of the open $3d$ shell. By this we do not mean the HF exchange polarization of the F^- ion by the $3d$ shell, a matter discussed elsewhere.[19] Instead, we mean the imbalance in covalent mixing caused by the imbalance in $3d$ exchange contributions to h_0 and, in turn, the various matrix elements appropriate to the determination of the covalent mixing. Although these exchange terms also contribute to the Ni^{2+} case, a more fundamental role is played by the intrinsic differences in the character of the antibonding-bonding pair covalency, on the one hand, and that of the unpaired bonding electrons on the other. Note that the unbalanced exchange terms cause the single determinant description of the Cr^{3+} cluster to be an improper eigenfunction, and of the Ni^{2+} cluster, to be a proper eigenfunction, of the spin operator (S^2). The Ni^{2+} description preserves its symmetry because of the presence of unpaired bonding orbitals of but one spin.[20]

V. THE CRYSTAL FIELD SPLITTING

We use the conventional definition of the crystal field splitting, namely,

$$10\,Dq \equiv \mathcal{E}[(t^{EA})^5(e^{EA})^3] - \mathcal{E}[(t^A)^6(e^A)^2], \quad (25)$$

where the latter is the energy of the ground state and the former is the energy of the excited state (hence the superscript E to designate this state) obtained by replacing the $t(xy)_\downarrow{}^A$ electron of the ground state by an $e(x^2-y^2)_\downarrow{}^A$ electron. If we make the common, but often unjustified,[21] assumption that all electrons but the promoted one maintain constant orbital behavior in the cluster states, we have

$$10\,Dq = \epsilon[e(x^2-y^2)_\downarrow{}^{EA}] - \epsilon[t(xy)_\downarrow{}^A], \quad (26)$$

where the ϵ's are the eigenvalues obtained with Eq. (6) for the one-electron Hamiltonian appropriate to the particular cluster state involved. To second order,

$$\epsilon_t{}^A = \langle\varphi_t|h|\varphi_t\rangle - 2S_t\langle\varphi_t|h|\chi_t\rangle - 2\gamma_t\langle\varphi_t|h|\chi_t\rangle + \{[S_t+\gamma_t]^2\langle\chi_t|h|\chi_t\rangle + [S_t{}^2-\gamma_t{}^2]\langle\varphi_t|h|\varphi_t\rangle\}, \quad (27)$$

which is obtained by multiplying Eq. (6) from the left by Ψ, reexpressing the Ψ's in terms of φ's and χ's, integrating and keeping all terms to second order. The first term of ϵ is simply the diagonal energy and includes the classic point charge crystal potential, Kleiner's correction[5] to such a potential, and ligand exchange terms; the second term is the first-order contribution to the overlap energy of the sort investigated by Tanabe and Sugano[7]; we then have the first-order term of the covalent contribution, which Sugano and Shulman concluded was important; and lastly, the second-order terms in overlap and covalency. The expression for $\epsilon_e{}^A$ is of the form of Eq. (27) except that there are first- and second-order overlap and covalent contributions from both the χ_s and χ_σ mixing, i.e.,

$$\epsilon_e{}^A = \langle\varphi_e|h|\varphi_e\rangle - \sum_{i=s,\sigma}\{2(S_i+\gamma_i)\langle\varphi_e|h|\chi_i\rangle + [S_i+\gamma_i]^2 \times\langle\chi_i|h|\chi_i\rangle + [S_i{}^2-\gamma_i{}^2]\langle\varphi_e|h|\varphi_e\rangle\}. \quad (28)$$

One may obtain alternative expressions[10] for the ϵ's, for example, by multiplying Eq. (16) by χ instead of Ψ, in which case ϵ^A is given by

$$\epsilon_t{}^A = \langle\varphi_t|h|\varphi_t\rangle - (S_t+\gamma_t)\langle\varphi_t|h|\chi_t\rangle + S_t(S_t+\gamma_t)\langle\varphi_t|h|\varphi_t\rangle. \quad (29)$$

Note the omission of the two multiplying the linear term.[22] Unlike Eq. (27), this relation only holds if Ψ is an eigenfunction of Eq. (6) (within our φ, χ subspace), in which case Eqs. (27) and (29) differ from one another in third and higher order in S and γ. This point is of some importance for we shall be evaluating Eq. (25) [and thus Eq. (27)] using noncovalent orbitals ($\gamma=0$).

We have already seen indications that the computed covalent mixing appropriate to a bonding-antibonding pair differs from that of an unpaired bonding electron. This suggests that the covalent mixing of the paired $t(xy)_\downarrow{}^B$ [and unpaired $e(x^2-y^2)_\downarrow{}^B$] orbital [$s$], appearing in the ground state, might very well undergo substantial changes on becoming an unpaired (and paired) electron in the excited state. We must then inspect the role of bonding orbital crystal field energies. The one-electron energy of a bonding t electron is, for example,

$$\epsilon_t{}^B = \langle\chi_t|h|\chi_t\rangle + 2\gamma_t\langle\varphi|h|\chi_t\rangle + \{\gamma_t{}^2\langle\varphi|h|\varphi\rangle - (2\gamma_tS_t+\gamma^2)\langle\chi_t|h|\chi_t\rangle\}. \quad (30)$$

[18] R. G. Shulman and K. Knox, Phys. Rev. Letters 4, 603 (1960).

[19] A. J. Freeman and R. E. Watson, J. Appl. Phys. 34, 1032S (1963); and (to be published).

[20] This observation is subject to the restriction that the Ψ's be constructed from a set of free ion φ's and ψ's having a single radial function per shell (although different φ_e and φ_t radial behavior is allowed).

[21] For example see, R. E. Watson, Phys. Rev. 118, 1036 (1960); in particular Tables V–VIII and related text.

[22] Their counterpart of the second line [e.g., in S&S III, Eq. (2.11)] shows a $(\varphi_e|h_0|\varphi_e)$ matrix element but we find $(\varphi_t|h_0|\varphi_t)$ more appropriate for this case.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The energy of a bonding-antibonding t orbital pair, Ω_t, in a crystal field is then

$$\begin{aligned}\Omega_t &= \epsilon_t{}^B+\epsilon_t{}^A-\langle\Psi_t{}^A\Psi_t{}^A|\Psi_t{}^B\Psi_t{}^B\rangle+\langle\Psi_t{}^A\Psi_t{}^B|\Psi_t{}^A\Psi_t{}^B\rangle\\ &= -S_t\langle\chi_t|h|\varphi_t\rangle+(1+S^2)\{\langle\varphi_t|h|\varphi_t\rangle+\langle\chi_t|h|\chi_t\rangle\\ &\qquad -\langle\varphi_t\varphi_t|\chi_t\chi_t\rangle+\langle\varphi_t\chi_t|\varphi_t\chi_t\rangle\}, \quad (31)\end{aligned}$$

where

$$\langle\Psi_A\Psi_B|\Psi_C\Psi_D\rangle \equiv \int\int d\tau_1 d\tau_2 \Psi_A{}^*(\mathbf{r}_1)\Psi_B(\mathbf{r}_1)\frac{1}{r_{12}}\Psi_C(\mathbf{r}_2)\Psi_D{}^*(\mathbf{r}_2). \quad (32)$$

The two-electron terms of the first line of (31) account for the fact that the interelectronic interaction between the pair of electrons has been counted twice when summing ϵ^B plus ϵ^A. An equivalent expression is obtainable for the e bonding electron; its derivation is but slightly complicated by the simultaneous presence of s and σ mixing.

In Eq. (31) as in Eqs. (16) and (17) we see the effect of the antibonding-bonding cancellation. The covalency of a bonding-antibonding pair plays no role in the cluster's energy[15] and hence in the crystal field splitting. Of course these pairs still make diagonal and overlap contributions to both the crystal field energy and to the Hamiltonian.

It is clear that only the unpaired bonding electrons make covalent contributions to the energies of Eq. (25). These total energies ($\mathcal{E}$) can be evaluated by a simple summation over one-electron ϵ's, providing that we account for the fact that this procedure introduces each interelectronic interaction twice by subtracting interelectronic Coulomb and exchange terms [as we did in Eq. (31)]. Here, we wish to concentrate on those ϵ's which contribute to a nonzero 10 Dq, namely all unpaired bonding electrons (which contribute covalent terms) and the bonding-antibonding pairs involved in the transition (which make diagonal and overlap contributions). The resulting $\mathcal{E}$'s are

$$\mathcal{E}[(t^A)^6(e^A)^2]=\epsilon_{xy\downarrow}{}^A+\epsilon_{xy\downarrow}{}^B+\epsilon_{s,v\downarrow}{}^B+\epsilon_{s,u\downarrow}{}^B+\epsilon_{\sigma,v\downarrow}{}^B+\epsilon_{\sigma,u\downarrow}{}^B+\sum_{i>j}[\langle\Psi_i\Psi_i|\Psi_j\Psi_j\rangle-\langle\Psi_i\Psi_j|\Psi_i\Psi_j\rangle]+\mathcal{E}_c \quad (33)$$

and

$$\mathcal{E}[(t^{EA})^5(e^{EA})^3]=\epsilon_{v\downarrow}{}^{EA}+\epsilon_{s,v\downarrow}{}^{EB}+\epsilon_{\sigma,v\downarrow}{}^{EB}+\epsilon_{xy\downarrow}{}^{EB}+\epsilon_{s,u\downarrow}{}^{EB}+\epsilon_{\sigma,u\downarrow}{}^{EB}+\sum_{i>j}[\langle\Psi_i\Psi_j|\Psi_j\Psi_j\rangle-\langle\Psi_i\Psi_j|\Psi_i\Psi_j\rangle]+\mathcal{E}_c, \quad (34)$$

where $\mathcal{E}_c$ is common to both energies and the sums go over all pairs of orbitals represented by the ϵ's. The subscripts v and u designate the (x^2-y^2) and $(3z^2-r^2)$ orbitals of e symmetry, respectively. The v orbital is involved in the crystal field transition ($t_{xy}{}^A \rightarrow e_{x^2-y^2}{}^A$) and the u orbitals are the "odd" unpaired bonding orbitals common to both crystal field states. Evaluating Eq. (25), rearranging terms and affecting appropriate cancellations, we have,

$$\begin{aligned}10Dq=&\{\langle\phi_e|h^E-V_{xy}{}^{\gamma E}-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}|\phi_e\rangle-\langle\phi_t|h-V_{s,v}{}^{\gamma}-V_{\sigma,v}{}^{\gamma}-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}|\phi_t\rangle\}\\ &+\sum_{\substack{i=s,j=\sigma\\ i=\sigma,j=s}}\{S_i{}^2[\langle\chi_{i,v}|h^E-V_{xy}{}^{\gamma E}-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}-\tfrac{1}{2}V_{j,v}{}^S|\chi_{i,v}\rangle+\langle\phi_v|h^E-V_{xy}{}^{\gamma E}-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}-\tfrac{1}{2}V_{j,v}{}^S|\phi_v\rangle]\\ &-2S_i\langle\chi_{i,v}|h^E-V_{xy}{}^{\gamma E}-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}-\tfrac{1}{2}V_{j,v}{}^S|\phi_v\rangle\}-\{S_\pi{}^2[\langle\chi_\pi|h-V_{s,v}{}^{\gamma}-V_{\sigma,v}{}^{\gamma}-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}|\chi_\pi\rangle\\ &+\langle\phi_{xy}|h-V_{s,v}{}^{\gamma}-V_{\sigma,v}{}^{\gamma}-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}|\phi_{xy}\rangle]-2S_\pi\langle\chi_\pi|h-V_{s,v}{}^{\gamma}-V_{\sigma,v}{}^{\gamma}-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}|\phi_{xy}\rangle\}\\ &-\{(\gamma_{xy\downarrow}{}^{EB})^2[\langle\chi_\pi|h^E-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}|\chi_\pi\rangle-\langle\phi_{xy}|h^E-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}|\phi_{xy}\rangle]\\ &-2\gamma_{xy\downarrow}{}^{EB}[\langle\chi_\pi|h^E-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}|\phi_{xy}\rangle-S_\pi\langle\chi_\pi|h^E-V_{s,u}{}^{\gamma E}-V_{\sigma,u}{}^{\gamma E}|\chi_\pi\rangle]\}\\ &+\sum_{\substack{i=s,j=\sigma\\ i=\sigma,j=s}}\{(\gamma_{i,v\downarrow}{}^B)[\langle\chi_{i,v}|h-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}-\tfrac{1}{2}V_{j,v}{}^{\gamma}|\chi_{i,v}\rangle-\langle\phi_v|h-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}-\tfrac{1}{2}V_{j,v}{}^{\gamma}|\phi_v\rangle]\\ &-2\gamma_{i,v\downarrow}{}^B[\langle\chi_{i,v}|h-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}-\tfrac{1}{2}V_{j,v}{}^{\gamma}|\phi_v\rangle-S_i\langle\chi_{i,v}|h-V_{s,u}{}^{\gamma}-V_{\sigma,u}{}^{\gamma}-\tfrac{1}{2}V_{j,v}|\chi_{i,v}\rangle]\}\\ &+\sum_{\substack{i=s,j=\sigma\\ i=\sigma,j=s}}\{(\gamma_{i,u\downarrow}{}^B)^2[\langle\chi_{i,u}|h-\tfrac{1}{2}V_{j,u}{}^{\gamma}|\chi_{i,u}\rangle-\langle\phi_u|h-\tfrac{1}{2}V_{j,u}{}^{\gamma}|\phi_u\rangle]-2\gamma_{i,u\downarrow}{}^B[\langle\chi_{i,u}|h-\tfrac{1}{2}V_{j,u}{}^{\gamma}|\chi_{i,u}\rangle\\ &-S_i\langle\chi_{i,u}|h-\tfrac{1}{2}V_{j,u}{}^{\gamma}|\phi_u\rangle]\}-\sum_{\substack{i=s,j=\sigma\\ i=\sigma,j=s}}\{(\gamma_{i,u\downarrow}{}^{EB})[\langle\chi_{i,u}|h^E-\tfrac{1}{2}V_{j,u}{}^{\gamma E}|\chi_{i,u}\rangle-\langle\phi_u|h^E-\tfrac{1}{2}V_{j,u}{}^{\gamma E}|\phi_u\rangle]\\ &-2\gamma_{i,u\downarrow}{}^{EB}[\langle\chi_{i,u}|h^E-\tfrac{1}{2}V_{j,u}{}^{\gamma E}|\chi_{i,u}\rangle-S_i\langle\chi_{i,u}|h^E-\tfrac{1}{2}V_{j,u}{}^{\gamma E}|\phi_u\rangle]\}, \quad (35)\end{aligned}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where each sum has but two terms ($i=s$ and $j=\sigma$ and vice versa) and where

$$V_i{}^S=S_i{}^2\left[\int\varphi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\varphi_i(r_2)d\tau_2+\int\chi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_i(r_2)d\tau_2\right]-2S_i\int\varphi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_i(r_2)d\tau_2, \tag{36}$$

$$V_i{}^\gamma=+2\gamma_i\left[\int\varphi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_i(r_2)d\tau_2-S_i\int\chi_i(r_2)\frac{1-P_{12}}{r_{12}}\chi_i(r_2)d\tau_2\right]$$
$$+\gamma^2\left[\int\varphi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\varphi_i(r_2)d\tau_2-\int\chi_i{}^*(r_2)\frac{1-P_{12}}{r_{12}}\chi_i(r_2)d\tau_2\right], \tag{37}$$

and h^E is the one-electron Hamiltonian defined for the excited state. The $\langle|V|\rangle$ integrals in Eq. (35) are the sole remaining interelectronic contributions from Eqs. (33) and (34), all others have cancelled. The first line of Eq. (35) consists of the diagonal crystal field terms; the next two lines are antibonding orbital overlap contributions; lines 4 and 5 are covalent contributions from the unpaired bonding t and e_v orbitals involved in the transition[23]; and the last two summations are covalent terms associated with the odd unpaired e_u bonding electrons. Within the cluster approximation Eq. (35) is an exact expression, to second order, for 10 Dq subject to the following three conditions:

(1) The Hamiltonian must be self-consistently defined for the set of Ψ's actually used, in which case Eq. (35) is correct independent of whether the Ψ's are, or are not, eigenfunctions of Eq. (6).

(2) The radial behavior of the φ and ψ atomic orbitals (not necessarily equal to that of free ion H-F functions), must be the same for the two crystal field states, otherwise there are additional contributions to 10 Dq.

(3) It is assumed that the ligand 1s shells do not enter into the bonding. Otherwise, there are additional covalent contributions from unpaired bonding $\Psi_{1s}{}^B$ orbitals. Such contributions are straightforwardly added to Eq. (35).

As already noted, the last four lines of Eq. (35) arise from the odd unpaired $\Psi_{i,u}$ bonding orbitals common to both crystal field states. These and the $V_{i,u}{}^\gamma$ contributions to the other lines sum to zero if an excited state parameter $\gamma_{i,u}{}^{EB}$ (and its associated matrix elements), equals its ground state, $\gamma_{i,u}{}^B$, counterparts. Assuming this equality to hold, Eq. (35) becomes more simply,

$$\begin{aligned}10\,Dq=&\{\langle\varphi_e|h^E|\varphi_e\rangle-\langle\varphi_t|h|\varphi_t\rangle\}+\{S_s{}^2[\langle\chi_s|h^E|\chi_s\rangle+\langle\varphi_e|h^E|\varphi_e\rangle]-2S_s\langle\chi_s|h^E|\varphi_e\rangle\}\\&+\{S_\sigma{}^2[\langle\chi_\sigma|h^E|\chi_\sigma\rangle+\langle\varphi_e|h^E|\varphi_e\rangle]-2S_\sigma\langle\chi_\sigma|h^E|\varphi_e\rangle\}-\{S_\pi{}^2[\langle\chi_\pi|h|\chi_\pi\rangle+\langle\varphi_t|h|\varphi_t\rangle]-2S_\pi\langle\chi_\pi|h|\varphi_t\rangle\}\\&-\{(\gamma_{\pi\downarrow}{}^{EB})^2[\langle\chi_\pi|h^E|\chi_\pi\rangle-\langle\varphi_t|h^E|\varphi_t\rangle]-2\gamma_{\pi\downarrow}{}^{EB}\langle\chi_\pi|h^E|\varphi_t\rangle+2\gamma_{\pi\downarrow}{}^{EB}S_\pi\langle\chi_\pi|h^E|\chi_\pi\rangle\}\\&+\{(\gamma_{s\downarrow}{}^B)^2[\langle\chi_s|h|\chi_s\rangle-\langle\varphi_e|h|\varphi_e\rangle]-2\gamma_{s\downarrow}{}^B\langle\chi_s|h|\varphi_e\rangle+2\gamma_{s\downarrow}{}^BS_s\langle\chi_s|h|\chi_s\rangle\}\\&+\{(\gamma_{\sigma\downarrow}{}^B)^2[\langle\chi_\sigma|h|\chi_\sigma\rangle-\langle\varphi_e|h|\varphi_e\rangle]-2\gamma_{\sigma\downarrow}{}^B\langle\chi_\sigma|h|\varphi_e\rangle+2\gamma_{\sigma\downarrow}{}^BS_\sigma\langle\chi_\sigma|h|\chi_\sigma\rangle\}\\&+\langle|V_{xy}{}^{\gamma E}|\rangle+\langle|V_e{}^\gamma|\rangle+\langle|V_{i,v}{}^S|\rangle.\end{aligned} \tag{38}$$

The three $\langle|V|\rangle$ terms have not been written out explicitly, in anticipation of the fact that we will evaluate this equation in terms of the approximate Hamiltonian, h_0, which does not contain V_s and V_γ terms. Hence, the $\langle|V|\rangle$ corrections of Eqs. (35) and (38) need not be made since these interelectronic effects were not counted twice in our sum over ϵ's. We will thus omit the last line when evaluating Eq. (38) with h_0. Without these terms, Eq. (38) would be identical with the traditional one given in terms of antibonding covalency, provided that (1) the ground-state unpaired $\gamma_{s\downarrow}{}^B$ and $\gamma_{\sigma\downarrow}{}^B$ values and their associated matrix elements are identical with those of the excited state $e_{s\downarrow}{}^{EA}$ and $e_{\sigma\downarrow}{}^{EA}$ electrons and (2) a similar equality holds for the unpaired $p_{\pi\downarrow}{}^{EB}$ and $p_{\pi\downarrow}{}^A$ contribitions. The present calculations indicate that these requirements are not met.

Before closing this section, we should note that, subject to two requirements, the unpaired bonding e electron covalent mixing contributing to 10 Dq is exactly that which is appropriate to the transferred hyperfine interactions. This differs with the antibonding orbital picture where ground-state antibonding spin ↑ covalency contributes to the hyperfine interaction and excited state antibonding spin ↓ covalency contributes to 10 Dq. The two requirements are: (1) that the covalent mixing of the two ground-state unpaired e^B orbitals be identical, a requirement fulfilled by symmetry considerations alone; and (2) that the "odd" e^B orbitals maintain constant covalency in the two cluster states, for

[23] When treating the covalent mixing of three orbitals (e.g., the $e_\downarrow{}^A$, $e_{s\downarrow}{}^B$, $e_{\sigma\downarrow}{}^B$) only two of which (the $e_{s\downarrow}{}^B$ and $e_{\sigma\downarrow}{}^B$) are occupied, there are covalent contributions to a quantity such as 10Dq. These include, first, the γ_s and γ_σ terms of Eq. (35) and second, nonzero $\gamma_{\sigma s}$ terms [see Eqs. (3)–(5) and related text] caused by nonzero $s-\sigma$ mixing. The $\gamma_{\sigma s}$ contributions enter in the third and higher orders in γ and S and were therefore not listed in Eq. (35). In addition, there are contributions arising from any nonzero overlap of closed Ni^{2+} shells with ligands, but these appear in third (and higher) order in 10 Dq.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

otherwise we have Eq. (35). It is reasonable to assume that this last requirement holds in a lower order theory but that, at some point, we must revert to using Eq. (35) where these orbitals contribute to 10 Dq.

VI. THE EVALUATION OF THE COVALENT MIXING AND THE ROLE OF SELF-ENERGY TERMS

We are interested in obtaining the unpaired bonding orbital covalency which contributes to 10 Dq and to transferred hyperfine effects. Also, as noted in Sec. II, independent estimates of the mixing occurring for the members of an antibonding-bonding pair supply us with a test of the internal consistency of our theory. With this factor in mind, we will estimate the covalency of all the occupied bonding and antibonding orbitals in the ground cluster state and of selected ones in the excited state. Let us now consider the process of estimating γ.

As already discussed, we wish to solve the H-F equations [Eq. (6)] defined[24] for the cluster, in a two-function space consisting of φ_t and χ_t for the t mixing, and a three-function space, made up of φ_e, χ_s and χ_σ for the e mixing, i.e., we solve 2×2 and 3×3 secular equations, respectively, for our Hamiltonian h. For the 3×3 solution, we need the matrix elements $\langle\chi_s|h|\chi_\sigma\rangle$ which were not considered by Sugano and Shulman. As these are intimately associated with the cluster approximation and also include important three- (and four-) center integrals and thereby divert us from the purpose of this paper, we shall not obtain these in the present treatment, deferring this matter, along with others which have been brought up, for a future effort involving the proper Hamiltonian, h, not h_0. We will instead explore the consequences of following the traditional view of assuming $\chi_s-\chi_\sigma$ mixing and overlap to be zero valued (i.e., $\gamma_{s\sigma}=\gamma_{\sigma s}=0$). In this approximation we obtain the following relation for any one of the γ values appropriate to Eq. (3):

$$S_i\langle\chi_i|h|\chi_i\rangle-\langle\varphi_1|h|\chi_i\rangle+\gamma_i[\langle\chi_i|h|\chi_i\rangle-\langle\varphi_i|h|\varphi_i\rangle]+\gamma_i^2[\langle\chi_i|h|\varphi_i\rangle-S_i\langle\varphi_i|h|\varphi_i\rangle]=0. \quad (39)$$

There are two roots to this equation: The $|\gamma|<1$ root is appropriate to the bonding orbital; the other, with $|\gamma|>1$, is appropriate to its orthogonal antibonding partner. Upon dropping the quadratic term, we obtain the Sugano-Shulman relation for γ_i, namely[25] [S&S III, Eq. (2.13)],

$$\gamma_i=\frac{\langle\varphi_i|h|\chi_i\rangle-S_i\langle\chi_i|h|\chi_i\rangle}{\langle\chi_i|h|\chi_i\rangle-\langle\varphi_i|h|\varphi_i\rangle}. \quad (40)$$

They also give a relation for λ_π [S&S III, Eq. (2.12)]

$$\lambda_\pi=\frac{\langle\varphi_t|h|\chi_\pi\rangle-S_\pi\langle\varphi_t|h|\varphi_t\rangle}{\langle\chi_\pi|h|\chi_\pi\rangle-\langle\varphi_\pi|h|\varphi_\pi\rangle}, \quad (41)$$

which we find is closely associated with the inverse of the second root of Eq. (39). We will obtain γ values appropriate to both bonding or antibonding mixing. We expect the smaller root of Eq. (39) and the result obtained from Eq. (40) to differ negligibly since the γ's are small; this will prove to be the case.

As discussed in Sec. III, in computing the matrix elements appearing in Eqs. (38)–(40), our H-F one-electron Hamiltonian, h, will be replaced by the approximate Hamiltonian h_0. Now, when evaluating the matrix element of a one-electron H-F Hamiltonian, one often may or may not, as one wishes, include the interelectronic Coulomb and exchange potential terms due to the electron for which the matrix element is being evaluated (in this case Ψ_i) because these terms, being equal, cancel one another. The inclusion of such self-energy terms leads to a common Hamiltonian for all electrons in the system, an important and often exploited feature of self-consistent H-F theory.[26,27] If by any chance the cancellation does not occur, these terms *must* be omitted from the Hamiltonian. Such a cancellation does not occur for h_0, since it is not a self-consistently defined Hamiltonian, and some ambiguity arises concerning the appropriate form of the self-energy term to be subtracted. Consider the evaluation of a matrix element appropriate to an antibonding orbital $\Psi_i{}^A$. The exact self-energy contribution to the potential is

$$\int\Psi_i{}^{A*}(r_2)\frac{1-P_{12}}{r_{12}}\Psi_i{}^A(r_2)d\tau_2, \quad (42)$$

but in view of the fact that h_0 is defined in terms of the (zero overlap) ionic ($Ni-F_6$) cluster [i.e., the first line of Eq. (8)], one would be subtracting out terms which did not appear in our Hamiltonian if we subtracted out Eq. (42). A possible choice, *consistent* with using the interelectronic potential of h_0, is to omit the Coulomb and exchange potential terms associated with the parent $3d\varphi_i$ (i.e., the overlap ionic counterpart of $\Psi_i{}^A$), as was done by Sugano and Shulman in the course of estimating antibonding covalency. Likewise, the equivalent χ_i contributions would then be omitted when estimating bonding covalency. The use of an approximate Hamiltonian has led to this choice of an approximate self-energy term. We will consider the effect of going to the more exact term after

[24] Since the antibonding-bonding pair mixing does not contribute to the cluster's total energy, we cannot properly derive the H-F equation appropriate to that mixing for such a derivation must follow from application of the variation principle to the total energy. One can write a one-electron H-F Hamiltonian [Eq. (7)] by inspection and solve for the pair mixing. This we will do but we must remember that the solution has a mathematical but not a physical meaning.

[25] Note the typographical error in the sign of the $S\langle\chi|h|\chi\rangle$ term in Eq. (2.13) of S&S III.

[26] For example, see A. J. Freeman and R. E. Watson, *Treatise on Magnetism*, edited by G. Rado and H. Suhl (Academic Press Inc., New York, to be published).

[27] Providing that we do not constrain spin pairs of bonding or antibonding orbitals to have the same covalent mixing. Such a requirement would involve us with the nonorthogonality associated with the conventional Hartree-Fock theory (Ref. 26).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

first considering the implications of this (zero-overlap) ionic approximation.

The question arises, as to whether the p_π bonding and antibonding matrix elements, for example, evaluated with different self-energy terms will, on insertion into Eqs. (40) and (41), yield $\lambda_\pi{}^A$ and $\gamma_\pi{}^B$ values which satisfy the orthogonality relation [cf. Eq. (5)]. A failure to satisfy Eq. (5) may arise from approximations in our Hamiltonian, in the treatment of the self-energy terms or in evaluating matrix elements; but, whatever its source, this failure implies a breakdown of theory. The matrix elements, for bonding and antibonding orbitals, are best compared by writing them for the proper ionic Hamiltonian, h_0 [the first line of Eq. (8) minus the appropriate φ or χ self-energy terms] and for a Hamiltonian, h_0', without the self-energy subtraction [literally the first line of Eq. (8)]. (Up to this point, we have not considered the role of self-energy terms in discussing the h_0 Hamiltonian. We here redefine h_0 to include the above self-energy term so as to conform to existing usage.) Upon doing this, we have the relations

$$\begin{aligned}
&\langle\varphi_i|h_0'|\chi_i\rangle^A-\langle\varphi_i|h_0|\chi_i\rangle^A=\langle\varphi_i\varphi_i|\varphi_i\chi_i\rangle-\langle\varphi_i\varphi_i|\varphi_i\chi_i\rangle,\\
&\langle\varphi_i|h_0'|\chi_i\rangle^B-\langle\varphi_i|h_0|\chi_i\rangle^B=\langle\chi_i\chi_i|\chi_i\varphi_i\rangle-\langle\chi_i\chi_i|\chi_i\varphi_i\rangle,\\
&\langle\varphi_i|h_0'|\varphi_i\rangle^A-\langle\varphi_i|h_0|\varphi_i\rangle^A=\langle\varphi_i\varphi_i|\varphi_i\varphi_i\rangle-\langle\varphi_i\varphi_i|\varphi_i\varphi_i\rangle,\\
&\langle\chi_i|h_0'|\chi_i\rangle^B-\langle\chi_i|h_0|\chi_i\rangle^B=\langle\chi_i\chi_i|\chi_i\chi_i\rangle-\langle\chi_i\chi_i|\chi_i\chi_i\rangle.
\end{aligned}\tag{43}$$

The integrals on the right-hand side of Eq. (43), coming from the Coulomb and exchange self-energy terms, obviously cancel, indicating that these matrix elements are independent of whether such terms are included in the Hamiltonian. However, we also have

$$\begin{aligned}
&\langle\varphi_i|h_0'|\varphi_i\rangle^B-\langle\varphi_i|h_0|\varphi_i\rangle^B=\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle,\\
&\langle\chi_i|h_0'|\chi_i\rangle^A-\langle\chi_i|h_0|\chi_i\rangle^A=\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle,
\end{aligned}\tag{44}$$

and here the cancellation of terms does not occur. Noting that the bonding and antibonding matrix elements of h_0' are, by definition, identical, gives us the following relationships between bonding and antibonding h_0 behavior for a bonding-antibonding pair:

$$\begin{aligned}
&\langle\varphi_i|h_0|\chi_i\rangle^A=\langle\varphi_i|h_0|\chi_i\rangle^B,\\
&\langle\varphi_i|h_0|\varphi_i\rangle^A=\langle\varphi_i|h_0|\varphi_i\rangle^B+\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle,\\
&\langle\chi_i|h_0|\chi_i\rangle^A=\langle\chi_i|h_0|\chi_i\rangle^B-\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle+\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle.
\end{aligned}\tag{45}$$

The two-electron integrals are most definitely nonzero.[28] But, what is more important, we will see that the $\langle|h_0|\rangle^A$'s, on insertion into, say, Eq. (39) yield a different γ than do the $\langle|h_0|\rangle^B$'s. This means that the theory, with its present approximations, fails to yield the required bonding-antibonding pair orthogonality. We will see this failure to be severe.

Not only does the theory suffer this internal inconsistency but it is difficult to ascertain which of the approximations is most at fault. Use of the full Hamiltonian of Eq. (8), the exact self-energy term [Eq. (42)], and extreme care in the evaluation of matrix elements would resolve this problem but would also be beyond the scope of this paper. We do wish to inspect one feature of the theory here, namely, given a Hamiltonian (in this case h_0'), what part do the self-energy terms play in the behavior of individual matrix elements and, in turn, on the resulting estimates of covalency. Such an investigation will further indicate the nature of the computed covalency and its sensitivity to computational details (including assumed covalency) while enabling us to avoid the serious numerical problems associated with a proper treatment with the proper covalent Hamiltonian. We believe this to justify a partial handling of the problem.

Let us define a Hamiltonian, H, which is equal to h_0' minus self-energy terms defined for the covalent Ψ_i's in Eq. (42). One may then obtain equations similar to (43) or (44) such as

$$\begin{aligned}
&\langle\varphi_i|h_0'|\chi_i{}^A\rangle-\langle\varphi_i|H|\chi\rangle^A\\
&\quad=N_i\{\langle\varphi_i\varphi_i|\varphi_i\chi_i\rangle-\langle\varphi_i\varphi_i|\varphi_i\chi_i\rangle-2\lambda_i\langle\varphi\chi|\varphi\chi\rangle\\
&\qquad+\lambda_i\langle\varphi\chi|\varphi\chi\rangle+\lambda_i\langle\varphi\varphi|\chi\chi\rangle+\lambda_i{}^2\langle\chi\chi|\varphi\chi\rangle\\
&\qquad\qquad-\lambda_i{}^2\langle\chi\varphi|\chi\chi\rangle\}\\
&\quad=-\lambda_i\{\langle\varphi\chi|\varphi\chi\rangle-\langle\varphi\varphi|\chi\chi\rangle\}+0(\lambda^3)
\end{aligned}\tag{46}$$

or

$$\begin{aligned}
\langle\varphi_i|H|\chi_i\rangle^A=\langle\varphi_i|h_0|\chi_i\rangle^A-\lambda_i\{&\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle\\
&-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\}.
\end{aligned}\tag{47}$$

For the other matrix elements we obtain

$$\begin{aligned}
&\langle\varphi_i|H|\chi_i\rangle^B=\langle\varphi_i|h_0|\chi_i\rangle^B+\gamma_i\{\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle\\
&\qquad\qquad-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\},\\
&\langle\varphi_i|H|\varphi_i\rangle^A=\langle\varphi_i|h_0|\varphi_i\rangle^A-\lambda^2\{\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle\\
&\qquad\qquad-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\},\\
&\langle\chi_i|H|\chi_i\rangle^A=\langle\chi_i|h_0|\chi_i\rangle^A+[1-N_i{}^2]\\
&\qquad\qquad\times\{\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\},\\
&\langle\varphi_i|H|\varphi_i\rangle^B=\langle\varphi_i|h_0|\varphi_i\rangle^B+[1-(N_i')^2]\\
&\qquad\qquad\times\{\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\},\\
&\langle\chi_i|H|\chi_i\rangle^B=\langle\chi_i|h_0|\chi_i\rangle^B-\gamma^2\{\langle\varphi_i\varphi_i|\chi_i\chi_i\rangle\\
&\qquad\qquad-\langle\varphi_i\chi_i|\varphi_i\chi_i\rangle\},
\end{aligned}\tag{48}$$

where terms of order λ^3, γ^3, or higher have been dropped. We will report solutions of Eq. (39) utilizing the linear covalent contributions to the off-diagonal $\langle\varphi|H|\chi\rangle^A$ and $\langle\varphi|H|\chi\rangle^B$ matrix elements (omitting second and

[28] The fact that self-energy terms cause $\langle|h|\rangle^B$ matrix elements to differ from their $\langle|h|\rangle^A$ counterparts is, in itself, a necessary but not sufficient condition for a computed γ^A to differ from γ^B. A computation with the exact one-electron Hamiltonian and with the exact self-energy terms [Eq. (42)] will have $\langle|h|\rangle^B\neq\langle|h|\rangle^A$ for the $\langle\chi|h|\chi\rangle$, $\langle\chi|h|\varphi\rangle$ and $\langle\varphi|h|\varphi\rangle$ matrix elements, one will obtain identical γ values. In the present case, the differences, in $\langle\varphi|h_0|\chi\rangle$ matrix elements will be seen to cause severe differences between a γ^A and its partner γ^B.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

higher order corrections). We will see the computed covalency to be sensitive to the estimates of covalency used in evaluating these matrix elements.

We have so far considered the differences which do, but should not, occur in the estimates appropriate to members of a bonding-antibonding pair. Let us now consider how these do (and should) differ with that appropriate to an unpaired bonding orbital. When obtaining the $\langle\varphi_i|h|\varphi_i\rangle$ and $\langle\varphi_i|h|\chi_i\rangle$ matrix elements for *either* member of a bonding-antibonding pair, the Coulomb and exchange contributions of φ_i to V_m [see Eq. (9)] cancel after the manner of the right-hand sides of Eqs. (43). The Coulomb repulsions of the *seven* other Ni $3d$ electrons remain, and the potential contributed by the Ni ion to these matrix elements is that characteristic of Ni^{2+}. However, no such cancellation occurs when evaluating the same matrix elements for an unpaired bonding orbital because the φ_i involved in the bonding is not one of the *eight* contributing to the Ni potential. Eight $3d$ electrons make Coulomb contributions in this case, causing the potential to be characteristic of Ni^{+}, not of Ni^{2+}. [Such a $3d$ electron Coulomb contribution to $\langle\varphi|h|\varphi\rangle$ which is $\sim$1 a.u., is to be compared with (1) the denominators of Eq. (40) which are as small as 0.25 a.u., (2) the $\langle\varphi\varphi|\chi\chi\rangle$ Coulomb self-energy terms of Eqs. (44) and (45), which are $\sim$0.3 a.u. and (3) $3d-3d$ exchange interactions, which are the order of 0.1 a.u.] It must be emphasized that this difference occurs whether we evaluate matrix elements with h_0 or with the exact self-consistent Hamiltonian h. In other words, such an effect on the current results involving h_0, is indicative of what actually occurs for the exact self-consistent LCAO treatment of the cluster.

From our experience with the potential terms of Sec. III, we should have anticipated the differences in the Ni potentials appropriate to the various $\langle\varphi|h|\varphi\rangle$ and $\langle\varphi|h|\chi\rangle$ matrix elements. The fact that these differences occur is a key to the intrinsically different nature of the covalency of a bonding-antibonding pair on one hand and an unpaired bonding orbital on the other. The magnitude (1 a.u.) of the difference in the $\langle\varphi|h|\varphi\rangle$ matrix elements has observable quantitive repercussions on one's theoretical estimates and makes it imperative that the pertinent covalent mixing be investigated.

VII. THE MATRIX ELEMENTS

In this section we evaluate the h_0 matrix elements appropriate to the various bonding and antibonding cases of interest, and examine the effects due to the net $3d$ spin on the Ni^{2+} ion. Wherever possible, we will rely on the one- and two-electron integrals already obtained in S&S III and the same approximations have been made when evaluating the V_L matrix elements. The additional required integrals have been evaluated with conventional (or restricted) analytic H-F (hence-forth denoted as RHF) functions for[29] Ni^{2+} and[30] F^{-}. However, no additional three- and four-center integrals were obtained for this work although they are required in a proper treatment of the theory. The Switendick-Corbato IBM 7090 program[31] was used for such two-center integrals as arose.

TABLE I. The one-electron energies (Ref. 32) of the RHF $3d$ orbitals, for the $^3F(M_L=L,\ M_s=S)$ state of Ni^{2+}, as defined by Eq. (52). In these tables, the appropriate m_s value is denoted by the arrow (↑ and ↓ denote $m_s=\frac{1}{2}$ and $-\frac{1}{2}$, respectively). Also included is the ϵ_{RHF} value. All quantities are in atomic units.

ϵ_{3d}	$m_s=$	↑	↓	
$m_l=2$		−1.46199	−1.39472	
1		−1.44377	−1.34007	
0		−1.41399	−1.34007	
−1		−1.44377	−0.38057	unoccupied
−2		−1.46199	−0.38057	orbitals
$\epsilon_{RHF}=-1.41254$				

A. $\langle\varphi_i|h_0|\varphi_i\rangle$

It is in this matrix element that $3d$ shell exchange and Ni^{+} versus Ni^{2+} potential repercussions are most important. We are constructing our molecular orbitals from free ion H-F orbitals and this suggests that we make use of the familiar[25] one-electron Hartree-Fock equation,

$$\mathcal{H}_{Ni}\varphi_i(\mathbf{r},\boldsymbol{\sigma})=\epsilon_i\varphi_i(\mathbf{r},\boldsymbol{\sigma}), \tag{49}$$

when evaluating $\langle\varphi_i|h_0|\varphi_i\rangle$. Here φ_i is a one-electron spin orbital (i.e., a function of space and spin), ϵ_i is its H-F energy eigenvalue, and $\mathcal{H}_{Ni}$ is the one-electron H-F Hamiltonian for Ni^{2+}, consisting of kinetic, nuclear potential, and interelectronic Coulomb and exchange terms. With the self-energy terms included, $\mathcal{H}_{Ni}$ is independent[26] of which Ni^{2+} orbital it operates on and is related to h_0' by

$$h_0'=\mathcal{H}_{Ni}+V_L(r). \tag{50}$$

Since our $3d\varphi_i$'s are RHF functions this suggests that we simply have

$$\langle\varphi_i|h_0'|\varphi_i\rangle=\epsilon_{RHF}+\langle\varphi_i|V_L|\varphi_i\rangle, \tag{51}$$

where ϵ_{RHF} is the $3d$ eigenvalue. Unfortunately, this equation does not hold, for the RHF $Ni^{2+}\varphi$'s are not exact eigenfunctions of Eq. (49). RHF theory requires[26] a single radial function per shell. For an open-shell ion this is only obtained by constraints, for there exists no single radial function which, when inserted into the occupied φ_i's, will satisfy Eq. (49) for all electrons in a shell. The immediate implications of this can be seen by first inspecting Table I where the values[32] of

$$\epsilon_i\equiv\langle\varphi_i|\mathcal{H}_{Ni}|\varphi_i\rangle \tag{52}$$

[29] R. E. Watson, Technical Report No. 12, Solid-State and Molecular Theory Group, MIT, 1959 (unpublished).
[30] C. Sonnenschein (unpublished).
[31] A. C. Switendick and F. J. Corbato, Quarterly Progress Report No. 34, Solid-State and Molecular Theory Group, MIT, October 1959 (unpublished).
[32] See p. 229 of Ref. 29.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. One-electron energies for the cubic $3d$ orbitals of Ni^{2+} as defined by Eq. (52) for the configurations indicated in the text. All quantities are in atomic units.

$\epsilon_{t\uparrow}=-1.42742$	$\epsilon_{e\downarrow}=-0.38057$
$\epsilon_{t\downarrow}=-1.35829$	$\epsilon_{e\downarrow}{}^{E}=-1.35829$
$\epsilon_{e\uparrow}=-1.47226$	$\epsilon_{t\downarrow}{}^{E}=-0.38057$
$\epsilon_{\mathrm{RHF}}=-1.41254$	

for the various occupied and unoccupied $3d$ orbitals of the RHF Ni^{2+} ion are listed. Also included is ϵ_{RHF}, which was used in S&S III when evaluating Eq. (51) for $\langle\varphi_i|h_0|\varphi_i\rangle$. We see that the occupied orbital ϵ_i's vary by as much as 0.12 a.u. Energy denominators as small as 0.22 a.u. occur in Eqs. (40) and (41), suggesting that these variations can observably affect the evaluation of these equations. One notes that the smaller ϵ_i's occur for occupied $3d$ orbitals of minority spin as one would expect since exchange lowers an electron's energy. The unoccupied orbital ϵ's lie approximately 1 a.u. higher than the others and show (cf. the discussion in the preceding section) the Coulomb energy penalty of attempting to place a ninth $3d$ electron on the ion. This penalty, which the unpaired bonding orbital covalency must pay, greatly influences results.

One may rigorously use Eq. (51) providing one inserts the ϵ_i, defined by Eq. (52) for the orbital *and* multielectron Ni state of interest. The ϵ_i's of interest are readily obtained with the available[29] integrals and are listed in Table II. The resulting $\langle\varphi_i|h_0|\varphi_i\rangle$'s evaluated using Eq. (51), with self-energy terms properly accounted for and utilizing the S&S III approximate values for the $\langle\varphi_i|V_L|\varphi_i\rangle$'s are given in Table III. Also listed for comparison are the matrix elements as obtained by Sugano and Shulman; these are $\langle\varphi_i|h_0|\varphi_i\rangle^A$ values.

B. $\langle\chi_i|h_0|\varphi_i\rangle$

If one assumes Eq. (49) to hold, then these matrix elements are obtained by first evaluating

$$\langle\chi_i|h_0'|\varphi_i\rangle=\langle\chi_i|\mathcal{H}_{\mathrm{Ni}}|\varphi_i\rangle+\langle\chi_i|V_L|\varphi_i\rangle$$
$$=\epsilon_{\mathrm{RHF}}S_i+\langle\chi_i|V_L|\varphi_i\rangle. \quad (53)$$

TABLE III. The matrix elements $\langle\varphi|h_0|\varphi\rangle$ for bonding (B), antibonding (A), and excited configuration (E) covalent mixing. All quantities are in atomic units.

		S&S III values			
$t_{\uparrow}{}^{A}$		0.1216	(0.1365)	$t_{\uparrow}{}^{B}$	−0.1358
$t_{\downarrow}{}^{A}$		0.1907	(0.1365)	$t_{\downarrow}{}^{B}$	−0.0667
$e_{\uparrow}{}^{A}$	(s or p_σ)	0.0604	(0.1202)	$e_{\uparrow}{}^{B}(s)$	−0.2058
				$e_{\uparrow}{}^{B}(p_\sigma)$	−0.2178
$e_{\downarrow}{}^{EA}$		0.1744	(0.1202)	$e_{\downarrow}{}^{EB}(s)$	−0.0918
				$e_{\downarrow}{}^{EB}(p_\sigma)$	−0.1038
				$e_{\downarrow}{}^{B}(s)$	0.8859
				$e_{\downarrow}{}^{B}(p_\sigma)$	0.8739
				$t_{\downarrow}{}^{EB}$	0.9110

But since the φ_i's are not exact eigenfunctions of that equation, the $\langle\chi_i|\mathcal{H}_{\mathrm{Ni}}|\varphi_i\rangle$ contribution should not be estimated in this way. This contribution has, with one approximation, been evaluated explicitly and the results appear in Table IV. The one approximation, necessitated by numerical inaccuracies in the $1s$ integral, consists of omitting the Coulomb potential term due to the $1s^2$ shell, and compensating for this with a nuclear potential term for a nuclear charge two less than that of Ni. Otherwise, the Switendick-Corbato program was used to obtain all integrals occurring in the $\langle\chi_i|\mathcal{H}_{\mathrm{Ni}}|\varphi_i\rangle$'s. The $\epsilon_{\mathrm{RHF}}S_i$ products are included in Table IV and we see that these differ but slightly from the exact values, indicating that assuming Eq. (42) has less significant repercussions here than it does for $\langle\varphi|h|\varphi\rangle$ matrix elements.

TABLE IV. Values of $\langle\chi_i|\mathcal{H}_{\mathrm{Ni}}|\varphi_i\rangle$ and $\epsilon_{\mathrm{RHF}}S_i$ (with S_i taken from S&S III) for ground and excited (E) configuration covalent mixing. All quantities are in atomic units.

	$\langle\chi_i\|\mathcal{H}_{\mathrm{Ni}}\|\varphi_i\rangle$	$\epsilon_{\mathrm{RHF}}S_i$
$t_{\uparrow}$	−0.10528	−0.1067
$t_{\downarrow}$	−0.10406	−0.1067
$e_{\uparrow}(s)$	−0.11751	−0.1150
$e_{\uparrow}(p_\sigma)$	−0.15150	−0.1564
$e_{\downarrow}{}^{E}(s)$	−0.11751	−0.1150
$e_{\downarrow}{}^{E}(p_\sigma)$	−0.14695	−0.1564
$e_{\downarrow}(s)$	−0.08478	...
$e_{\downarrow}(p_\sigma)$	−0.09572	...
$t_{\downarrow}{}^{E}$	−0.0739	

Using the S&S III estimates of the $\langle\chi_i|V_L|\varphi_i\rangle$ contributions, the $\langle\chi_i|h_0|\varphi_i\rangle$ matrix elements are given in Table V. As we see from Eqs. (45), these matrix elements, unlike $\langle\chi|h_0|\chi\rangle$ and $\langle\varphi|h_0|\varphi\rangle$, are independent of whether we are dealing with bonding or antibonding orbitals. The matrix elements are sensitive to whether they are for an antibonding-bonding pair or for an unpaired bonding orbital [$e(s)_{\downarrow}$, $e(p\sigma)_{\downarrow}$, and $t_{\downarrow}{}^{E}$]. We see values for the latter to be appreciably smaller in magnitude.

TABLE V. The matrix elements $\langle\chi_i|h_0|\varphi_i\rangle$ for ground and excited (E) configuration covalent mixing.[a] All quantities are in atomic units.

	Present calculation	S&S III
$t_{\uparrow}$	−0.0428	(−0.0443)
$t_{\downarrow}$	−0.0416	
$e_{\uparrow}(s)$	−0.1185	(−0.1160)
$e_{\uparrow}(p_\sigma)$	−0.0937	(−0.0986)
$e_{\downarrow}{}^{E}(s)$	−0.1185	
$e_{\downarrow}{}^{E}(p_\sigma)$	−0.0891	
$e_{\downarrow}(s)$	−0.0858	
$e_{\downarrow}(p_\sigma)$	−0.0379	
$t_{\downarrow}{}^{E}$	−0.0115	

[a] These matrix elements are independent of whether we are dealing with bonding or antibonding effects (see text).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

C. $\langle \chi_i | h_0 | \chi_i \rangle$

The various values of this matrix element are readily obtained using the integrals appearing[33] in Table II of S&S III. The results appear in Table VI. Two features of the matrix element behavior should be noted. First, as expected, unlike the cases of $\langle \chi | h_0 | \varphi \rangle$ and $\langle \varphi | h_0 | \varphi \rangle$ matrix elements, the $\langle \chi | h_0 | \chi \rangle$ matrix elements appropriate to the unpaired bonding orbitals differ but trivially from the paired bonding orbital values. This small difference indicates that these matrix elements cannot compensate for the shifts, appearing in the unpaired bonding $\langle \varphi | h_0 | \chi \rangle$ and $\langle \varphi | h_0 | \varphi \rangle$ values, due to

TABLE VI. The matrix elements $\langle \chi_i | h_0 | \chi_i \rangle$ for bonding (B), antibonding (A) and excited configuration (E) covalent mixing. All quantities are in atomic units.

$t_\uparrow{}^A$	−0.0837		$t_\uparrow{}^B$	+0.1737
$t_\downarrow{}^A$	−0.0835	(−0.0835)	$t_\downarrow{}^B$	+0.1739
$e_\uparrow{}^A(s)$	−0.9965		$e_\uparrow{}^B(s)$	−0.7303
$e_\uparrow{}^A(p_\sigma)$	−0.1659		$e_\uparrow{}^B(p_\sigma)$	+0.1123
$e_\uparrow{}^{EA}(s)$	−0.9959	(−0.9959)	$e_\downarrow{}^{EB}(s)$	−0.7297
$e_\uparrow{}^{EA}(p_\sigma)$	−0.1628	(−0.1628)	$e_\downarrow{}^{EB}(p_\sigma)$	+0.1154
			$e_\downarrow{}^B(s)$	−0.7300
			$e_\downarrow{}^B(p_\sigma)$	+0.1162
			$t_\downarrow{}^{EB}$	+0.1753

the Ni^+ potential. Secondly, we see a sign reversal, similar to the one in Table II, on going from the antibonding to the bonding member of a p_π or p_σ pair, which is again due to the self-energy terms of Eqs. (45). Their effect is to reverse the sign of the denominator of Eq. (40) and of the linear term of Eq. (39), a feature which will be of great importance to us when discussing results.

D. The Overlap Integrals, S_i

As already noted, the present investigation has utilized analytic HF functions in its evaluation of integrals. Sugano and Shulman used the same Ni $3d$, but different F^-, orbitals. They used simple two-exponential analytic approximations to the F^- functions of Froese,[34] a choice which was necessitated by the extensive nature of the computations. S_i values which were obtained for the current set of functions, along with their S&S III counterparts, are listed in Table VII; the differences serve as a measure of the error, due to wave function uncertainty, that should be attached to such integrals as these. The S_i obtained with Sonnenschein's F^- orbitals will be used in the following section.

TABLE VII. Comparison of overlap integrals obtained by S&S III and in this calculation.

	This calculation	S&S III
S_t	0.07077	0.07557
$S_e(s)$	0.07977	0.08143
$S_e(p_\sigma)$	0.10499	0.11071

[33] Two additional integrals are needed. Their values, in the notation of the S&S III Appendix, are: $(d\sigma, p\pi^+ \| d\delta^+, p\pi^-) = -0.000255$ a.u. $(d\sigma, p\pi^+ \| p\pi^-, d\delta^+) = 0.000039$ a.u. The $\langle \chi | h_0 | \chi \rangle$ integrals are the most seriously affected by the (almost) complete omission (Ref. 10) of three and four-center integrals.

[34] C. Froese, Proc. Cambridge Phil. Soc. 53, 206 (1957).

E. $\langle | H | \rangle$

From the definitions of Eqs. (47) and (48), we see that these matrix elements are functions of covalency and as we will want them for varying γ and λ values, we list in Table VIII the terms

$$[\langle \varphi_i \varphi_i | \chi_i \chi_i \rangle - \langle \varphi_i \chi_i | \varphi_i \chi_i \rangle],$$

which are necessary for evaluating these equations and the corresponding terms of Eq. (45) (used in the evaluation of Tables III and VI). These were obtained from integrals appearing in S&S III, which only include two-center integral contributions. It should be noted that by basing these terms on two-center integrals alone, we are making ligand χ self-energy corrections [the first equation of (44)] which are strictly compatible with the V_L in our h_0 Hamiltonian. Here too, the inclusion of three-center contributions to these integrals may have appreciable quantitative effects on these terms. In the calculations which follow, we will keep only the linear covalent contributions occurring in the $\langle \varphi_i | H | \chi_i \rangle$ matrix elements; the other elements will then keep their h_0 values. In view of our decision not to use a proper covalent Hamiltonian in this paper, a more detailed treatment seems unjustified.

VIII. RESULTS

Equations (39) and (40) have been evaluated with the h_0 matrix elements of the preceding section for the various bonding and antibonding γ's. The results appear in Table IX. Only the γ's, which are less than 1 in magnitude, have been reported for Eq. (39); as stated earlier, the larger roots are redundant.

An important feature of the results in Table IX concerns the internal consistency of a theory relying on h_0. We see that the $\gamma_i{}^A$ and $\gamma_i{}^B$, of a p_σ or p_π antibonding-bonding pair, differ in sign while those for s bonding differ in magnitude. As we have already stressed, a $\gamma_i{}^A$ must equal its partner $\gamma_i{}^B$ if the orthogonality requirements of Eqs. (5) are to be realized. Any failure to meet these requirements implies a

TABLE VIII. The $[\langle \varphi_i \varphi_i | \chi_i \chi_i \rangle - \langle \varphi_i \chi_i | \varphi_i \chi_i \rangle]$ values necessary for obtaining the covalent contributions to the $\langle | H | \rangle$ matrix elements. All quantities are in atomic units.

i	$[\langle \varphi_i \varphi_i \vert \chi_i \chi_i \rangle - \langle \varphi_i \chi_i \vert \varphi_i \chi_i \rangle]$
$e(s)$	0.26618
$e(p_\sigma)$	0.27823
$t(p_\pi)$	0.25741

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE IX. The γ_i covalent mixing parameters, as estimated with Eqs. (39) and (40), for bonding (B), antibonding (A) and excited configuration (E) cases, and compared with the S&S III values.

		Eq. (40)	Eq. (39)	S&S III values			Eq. (40)	Eq. (39)
$t_\uparrow^A$		0.180	0.171		$t_\uparrow^B$		−0.178	−0.174
$t_\downarrow^A$	c	0.130	0.127	(0.1727)	$t_\downarrow^B$		−0.224	−0.217
$e_\uparrow^A(s)$	d	0.037	0.037		$e_\uparrow^B(s)$		0.114	0.112
$e_\uparrow^A(p_\sigma)$	d	0.337	0.298		$e_\uparrow^B(p_\sigma)$		−0.319	−0.287
$e_\downarrow^{EA}(s)$	c	0.033	0.033	(0.0313)	$e_\downarrow^{EB}(s)$		0.094	0.093
$e_\downarrow^{EA}(p_\sigma)$	c	0.214	0.201	(0.2848)	$e_\downarrow^{EB}(p_\sigma)$		−0.462	−0.404
					$e_\downarrow^B(s)$	a, b	0.0171	0.0169
					$e_\downarrow^B(p_\sigma)$	a, b	0.0661	0.0654
					$t_\downarrow^{EB}$	a	0.0325	0.0322

[a] Contribute to 10 Dq.
[b] Contribute to transferred hyperfine effects.
[c] Contribute to 10 Dq in the traditional antibonding model.
[d] Contribute to the transferred hyperfine effects in the traditional antibonding model.

breakdown of the theory. The deviations seen in Table IX lead to antibonding-bonding orthogonality integrals as large as 0.6 instead of the required zero. Thus, the severity of this orthogonality breakdown renders any quantity estimated with these γ's at best suspect.

The γ_p sign reversals of Table IX have been caused by the reversal in sign of the $[\langle\chi|h_0|\chi\rangle-\langle\varphi|h_0|\varphi\rangle]$ term of Eqs. (39)–(41), due to self-energy effects. We expect a similar sign reversal in this term if we go to a theory utilizing the exact self-consistent Hamiltonian (and the exact self-energy expressions). The only way in which this exact theory can then yield a common $\gamma_p{}^A$ and $\gamma_p{}^B$ will be if $\langle\varphi|h|\chi\rangle$ reverses sign as well (note that here we refer to a matrix element of h, not h_0). We will shortly see indications that this does indeed occur.

Three other features of the results of Table IX should be viewed before we leave them. First, Eq. (40) and the more exact (39) give results of similar character but occasionally differ by as much as 10%. Secondly, the difference in γ's, for pairs of orbitals differing only in spin, are not negligible. These do not affect our predictions for Ni^{2+} but such effects are of sufficient magnitude to observably affect the transferred hyperfine spectrum associated with an ion such as Cr^{3+} (cf. Sec. IV). Finally, we see the unpaired bonding orbital γ values to be appreciably smaller than their paired orbital antibonding counterparts, suggesting that one must not assume them equivalent and that one must utilize the unpaired bonding results when estimating 10 Dq or some other experimental parameter. The large uncertainty which one must attach to the antibonding γ's (due to the antibonding-bonding pair breakdown) makes this comparison tenuous. The results which we will now inspect, also indicate substantial differences between the γ^A and unpaired γ^B behavior.

We are not prepared in the present paper to attempt a calculation involving the full Hamiltonian, h, of Eq. (8), but we do wish to gain some idea of the sensitivity of the predictions to covalency. To do this we will go over to the covalent treatment of the self-energy correction, utilizing the H matrix elements of Eqs. (47) and (48) (actually we will only consider the linear contributions to the $\langle\varphi|H|\chi\rangle$'s). Computing γ_i's as a function of assumed γ^0 values (appearing in $\langle\varphi|H|\chi\rangle$), we obtain the results of Table X. We see that the bonding-antibonding pair results are quite sensitive to this covalent contribution and that the negative γ's turn positive with increasing covalency because of a sign reversal in the $\langle\varphi|H|\chi\rangle$ elements. Since we have not used the proper Hamiltonian, we cannot expect complete convergence of the p electron bonding-antibonding γ_i pairs. The table also suggests that, given the Hamiltonian, we have underestimated the bonding-antibonding covalency if we use the antibonding results of Table IX.

In contrast with the paired bonding behavior, the unpaired bonding γ's of Table X show a tendency of $\langle\varphi|H|\chi\rangle$ to reverse sign. The $(\langle\chi|h_0|\chi\rangle-\langle\varphi|h_0|\varphi\rangle)$ term does not have a negative sign (as do the paired bonding elements), and this causes the trend toward negative γ's. The tendency towards increased paired orbital γ's and decreased unpaired bonding γ's serves to accentuate the different nature of the two types of covalent mixing.

TABLE X. The covalent mixing γ_i, as estimated using Eq. (39) with $\langle\varphi_i|H|\chi_i\rangle$ matrix elements for bonding (B), antibonding (A), and excited configuration (E) cases as a function of assumed covalency.

	Assumed $\gamma_i{}^0$						
	0	0.05	0.1	0.2	0.3	0.4	0.5
$e_\uparrow^A(s)$	0.057	0.069	0.081	···	···	···	···
$e_\uparrow^B(s)$	0.112	0.088	0.064	···	···	···	···
$e_\uparrow^A(\sigma)$	0.38	···	0.44	0.50	0.55	0.59	0.62
$e_\uparrow^B(\sigma)$	−0.30	···	−0.23	−0.15	−0.07	0.01_7	0.10
$t_\uparrow^A$	0.24	···	0.34	0.42	0.48	0.53	0.58
$t_\uparrow^B$	−0.17	···	-0.09_4	−0.12	−0.07	0.15	0.22
$t_\downarrow^A$	0.19	···	0.27	0.33	0.39	0.44	0.49
$t_\downarrow^B$	−0.22	···	−0.12	−0.10	-0.09_8	0.19	0.28
$e_\downarrow^{EA}(\sigma)$	0.27	···	0.33	0.38	0.43	0.47	0.50
$e_\downarrow^{EB}(\sigma)$	−0.40	···	−0.31	−0.20	-0.08_5	0.04_6	0.16
$e_\downarrow^{EA}(s)$	0.051	0.062	0.073	···	···	···	···
$e_\downarrow^{EB}(s)$	0.093	0.073	0.053	···	···	···	···
$e_\downarrow^B(s)$	0.017	0.009	0.0008	−0.016	···	···	···
$e_\downarrow^B(\sigma)$	0.065	0.047	0.0295	−0.007	···	···	···
$t_\downarrow^{EB}$	0.032	0.015	0.0030	−0.033	···	···	···

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE XI. The covalent mixing parameters, γ_i, as estimated using Eq. (39) with $\langle\varphi_i|H|X_i\rangle$ matrix elements for the unpaired $t_\downarrow^{EB}$, $e_\downarrow^B(s)$, and $e_\downarrow^B(\sigma)$ orbitals for assumed γ_i^0 values.

	Input γ_i^0	Computed γ_i
$e_\downarrow^B(s)$	0.01	0.015_4
$e_\downarrow^B(\sigma)$	0.05	0.047
$t_\downarrow^{EB}$	0.025	0.0236

In Table XI we have reevaluated the unpaired γ's with input $\langle\varphi|H|\chi\rangle$ covalent contributions suggested by Table X. The results show internal consistency and are in crude agreement with the noncovalent estimates of Table IX. This is largely due to the fact that the covalency is small and hence the covalent self-energy repercussions are small (but not insignificant).

So far we have only considered the occupied orbitals of the cluster and have not used the alternate approach of dealing with the antibonding holes. As stated earlier, such an approach is exactly equivalent to the unpaired bonding method provided that one maintains various restrictions of the present paper (e.g., common radial behavior for the φ's and χ's appropriate to an antibonding-bonding pair on the one hand and to the unpaired bonding orbitals and antibonding holes on the other). In a less restricted treatment, the two approaches need not be equivalent but one is then involved with an unrestricted H-F theory and its associated difficulties.[26] Estimates of antibonding hole γ values involve *no* self-energy corrections [e.g., in Eqs. (44), (47), and (48)], and the results must equal (to second order) their unpaired bonding counterparts in order to maintain orthogonality requirements. Using Eq. (40) and the matrix elements and integrals of Tables III–VIII, we obtain values of γ equal to 0.014_7, 0.048, and 0.024 for the $e(s)$, $e(\sigma)$, and t^E orbitals, respectively. These are in remarkable agreement with the values listed in Table XI.

While we seem to have converged on "self-consistent" unpaired bonding values in Table XI, this by no means implies that they are in detailed agreement with what will be yielded by the more exact theory. On the other hand, they differ strikingly with those bonding-antibonding results obtained either by Sugano and Shulman or in Tables IX and X. This is due to the Ni^+-Ni^{2+} Coulomb cancellation differences and will remain in the exact self-consistent field cluster theory.

In order to facilitate the discussion of 10 Dq and transferred hyperfine effects, the γ values, appropriate to the unpaired bonding and antibonding models, have been extracted from Tables IX to XI and listed together in Table XII. The S&S III antibonding results are included for comparison, as are the γ^A values suggested by Table X, since these would be appropriate to 10 Dq and the transferred hyperfine effects if we had adhered to the traditional (but incorrect) antibonding model. Different γ^A values are reported for 10 Dq and for the transferred hyperfine effects as spin $\downarrow$ parameters are appropriate to the former and spin $\uparrow$ to the latter in the antibonding model. Computed 10 Dq or f_i values based on these γ^A will not be reported since this would be incorrect. We merely note that the use of these γ^A's would predict even stronger covalent effects than those appearing in S&S III.

A. Transferred Hyperfine Effects

Values of f_i were determined by using Eqs. (23) and (24) for the various sets of unpaired bonding γ_i's reported (and defined) in Table XII. The separate spin $\uparrow$ and spin $\downarrow$ contributions to the f_i values are listed in Table XIII along with the experimental[11] values of f_i. The experimental f_s value includes the so-called $1s-2s$ cross-term correction.[16,17] Under the antibonding column, we list the S&S III predictions and give their pure overlap contribution separately from the covalent terms [cf. Eq. (20)].

The predicted f_s values are not in good agreement with experiment; the S&S III antibonding prediction is in somewhat better, but by no means good, agreement. What agreement there exists for the unpaired bonding results, arises largely from the spin $\uparrow$ electron overlap term rather than from the spin $\downarrow$ covalent terms.

The experimental f_σ value lies between the S&S III antibonding prediction and those of the present calculations. The theory is again seen to be in poor agreement with experiment. Quite aside from the fact that the S&S III predictions involve the inappropriate antibonding γ's, we do not find them in marked numerical agreement with experiment. The present results consistently and appreciably underestimate both f_s and f_σ. We will see a similar, and perhaps related, low value for 10 Dq. Let's consider that case now.

B. Crystal Field Splitting Parameter

Our expression for 10 Dq [Eq. (38)] was derived for the exact self-consistent Hamiltonian h, but, as discussed earlier, we shall use h_0 matrix elements to estimate the diagonal, overlap, and covalent contributions. Before proceeding to this task, it is instructive to

TABLE XII. The γ values appropriate to the unpaired bonding and the antibonding models for 10Dq and the transferred hyperfine effects as obtained from Tables IX–XI. Included are the S&S III antibonding estimates.

Unpaired bonding				Antibonding	
A	*B*	*C*		*D*	*E*
Table IX Eq. (40)	Table IX Eq. (39)	Tables X, XI		S&S III	Table X
For transferred hyperfine effects					
$\gamma_{s\downarrow}{}^B$ 0.017_+	0.017_-	0.015	$\gamma_{s\uparrow}{}^A$	0.031	0.05–0.08
$\gamma_{\sigma\downarrow}{}^B$ 0.066	0.065	0.047	$\gamma_{\sigma\uparrow}{}^A$	0.285	0.4 –0.6
For 10Dq					
$\gamma_{s\downarrow}{}^B$ 0.017_+	0.017_-	0.015_4	$\gamma_{s\downarrow}{}^{EA}$	0.031	0.05–0.07
$\gamma_{\sigma\downarrow}{}^B$ 0.066	0.065	0.047	$\gamma_{\sigma\downarrow}{}^{EA}$	0.285	0.2 –0.5
$\gamma_{\pi\downarrow}{}^{EB}$ 0.032_5	0.032_2	0.024	$\gamma_{\pi\downarrow}{}^A$	0.173	0.2 –0.5

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE XIII. Comparison of unpaired bonding [cf. Eqs. (23) and (24)] and antibonding (S&S III) transferred hyperfine parameters, f_i, with experiment. The various calculations are defined as in Table XII.

	Unpaired bonding			Antibonding
f_s	A	B	C	S&S III
Spin ↑ overlap	=0.0021	0.0021	0.0021	Spin ↑ overlap =0.0022
Spin ↓ covalent	=0.0010	0.0010	0.0009	Spin ↑ covalent =0.0020
Total f_s	=0.0031	0.0031	0.0030	Total f_s =0.0042
Experiment	=0.0054			
f_σ				
Spin ↑ overlap	=0.0037	0.0037	0.0037	Spin ↑ overlap =0.0041
Spin ↓ covalent	=0.0060	0.0059	0.0040	Spin ↑ covalent =0.0480
Total f_σ	=0.0097	0.0096	0.0081	Total f_σ =0.0521
Experiment	=0.0378			

compare this estimate with the results appearing in S&S III. Such a comparison, however, cannot be done directly since their equation for 10 Dq [S&S III, Eq. (2.11)] differs with Eq. (38). For example, they have an e_s overlap term,

$$-S_s\langle\varphi_c|h|\chi_s\rangle+S_s^2\langle\varphi_c|h|\varphi_c\rangle \tag{54}$$

and a covalent contribution

$$-\gamma_s^{EA}\langle\varphi_c|h|\chi_c\rangle+S_s\gamma_s^{EA}\langle\varphi_c|h|\varphi_c\rangle. \tag{55}$$

Here, the second-order contributions appear in the "renormalization" term of their equation. Equations (54) and (55) obviously differ with lines 1 and 4 of Eq. (38) which are their counterparts here. Prior to comparing our results with S&S III, we must then ask whether the differing equations for 10 Dq in any way affect our observations. With this in mind, we will first evaluate Eq. (38) with the antibonding γ's and matrix elements of S&S III. [Equation (38) is readily converted for use in their antibonding model by replacing the values of γ and the matrix elements of the fifth line (for the t^{EB} electron) by the corresponding t^A values and similarly replacing the e_s^B and e_σ^B of the sixth and seventh lines by their counterparts e_s^{EA} and e_σ^{EA}, respectively.] The results of this calculation are given in Table XIV along with the values obtained by using Eq. (2.11) of S&S III, with and without second-order terms. Sugano and Shulman identified the linear terms, alone, as specific overlap or covalent contributions. With the inclusion of second-order terms, we see from Table XIV that the two equations yield the same *total* contribution from a particular electron but differ appreciably as to the relative roles of "covalent" and "overlap" effects. This implies that one equation or the other has incorrectly identified these contributions.

As already discussed, Eq. (38) was obtained from one-electron energy expressions [e.g., Eqs. (27), (28), and (30)] by the explicit evaluation of $\langle\Psi|h|\Psi\rangle$, a process which correctly yields the energy whether or not Ψ is an eigenfunction of h. The matrix elements of the resulting equations were then in no way manipulated. Equation (38) correctly yields a value of 10 Dq in the limit of zero covalency by the simple expedient of setting the γ values equal to zero: the terms which remain are considered to be diagonal or overlap contributions; those which are introduced in the process of allowing nonzero γ's are considered covalent effects. Thus, we conclude, Eq. (38) gives the proper breakdown of covalent and overlap effects.

Equation (2.11) of S&S III is based on one-electron energy expressions such as Eq. (29) which in turn was obtained from

$$\langle\varphi|h|\Psi\rangle/\langle\varphi|\Psi\rangle=\epsilon, \tag{56}$$

a relation which holds only if Ψ is an eigenfunction of the H-F equation [Eq. (6)], i.e., only for the *one* value of γ appropriate to that equation. One can alternatively obtain Eq. (29) by using Eq. (40) to remove a linear $-(\gamma+S)\langle\varphi|h|\chi\rangle$ term from Eq. (27). However, in doing this, one assumes a specific relation between γ, S, $\langle\varphi|h|\varphi\rangle$, $\langle\varphi|h|\chi\rangle$, and $\langle\chi|h|\chi\rangle$, namely, that γ be appropriate to the H-F solution. The resulting expression is simpler than Eq. (38), but it yields correct results only if the H-F γ value is inserted, and we can no longer correctly separate the covalent and overlap contributions by simple inspection, as we did for Eq. (38). From Table XIV we see that in S&S III the role of overlap effects was underestimated and that of covalency overestimated.

Computed unpaired bonding values for 10 Dq are listed in Table XV: Cases A and B (described earlier) have been merged since they yield identical results; the S&S III antibonding values[35] of Table XIV have again been included for comparison. All of these results have been obtained by using Eq. (38) and all indicate

TABLE XIV. The overlap and covalent contributions to 10 Dq obtained using Eq. (38) [modified as indicated in the text] and by using Eq. [S&S III, (2.11)] when evaluated with the matrix elements and γ^A's of S&S III. All quantities are in cm^{-1}.

		S&S III Eq. (2.11)	
	Eq. (38)	Linear terms only	Linear plus second-order terms
s contributions			
overlap	2870	2070	2250
covalent	240	790	855
total	3110		3105
$p\sigma$ contributions			
overlap	4680	2400	2720
covalent	5035	6170	7000
total	9715		9720
$p\pi$ contributions			
overlap	−1535	−730	−905
covalent	−1440	−1680	−2075
total	−2975		−2980
Total overlap	6015	3740	4065
Total covalent	3835	5280	5780
Total	9850	9020	9845

[35] The total antibonding value for 10Dq differs by 70 cm^{-1} with that appearing in S&S III. This occurs because we differ as to the matrix element appropriate to the second-order t_{2g} orbital contribution (see footnote 22).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE XV. The diagonal (as estimated in S&S III), overlap and covalent contributions to 10 Dq evaluated with Eq. (38) and the h_0 Hamiltonian for the unpaired bonding γ's of sets A–B and C, and for the antibonding γ's of S&S III. All quantities are in cm^{-1}.

	Case A–B	Case C		S&S III
Diagonal	−3570	−3570		−3570
Overlap				
$e_s{\downarrow}^{EA}$	3000	3000		2870
$e_\sigma{\downarrow}^{EA}$	4135	4135		4680
$t_\downarrow{}^{A}$	−1410	−1410		−1535
Total overlap	5725	5725		6015
Covalent				
Unpaired $e_s{\downarrow}^{B}$	105	100	$e_s{\downarrow}^{EA}$	240
Unpaired $e_\sigma{\downarrow}^{B}$	725	665	$e_\sigma{\downarrow}^{EA}$	5035
Unpaired $t_\downarrow{}^{EB}$	−170	−160	$t_\downarrow{}^{A}$	−1440
Total covalent	660	605		3835
10 Dq=	2815	2760		6280

Experimental 10 Dq=7250

that overlap effects dominate. The covalent contributions play a small role in the present calculations and the resulting 10 Dq values are observed to be substantially smaller than experiment.

Upon seeing such a disagreement with experiment, one might well inquire into the role played by uncertainties or errors in the γ values. From a given set of matrix elements (say, the unpaired bonding $\langle|h_0|\rangle$'s), an unpaired bonding γ, which is obtained variationally, gives the lowest one-electron (and in turn cluster) covalent crystal field energy obtainable for the orbital in question. Any deviation in the γ eigenvalues, for electrons appearing in the cluster ground state, will raise the ground-state energy, hence decreasing 10 Dq. Any similar deviation for a γ^E will raise the excited state energy, thereby increasing 10 Dq. If we utilize the unpaired bonding $\langle|h_0|\rangle$ integrals, any variation in the $\gamma_\sigma{}^B$ or $\gamma_s{}^B$, from the values of Table IX, will, therefore, decrease 10 Dq, while a variation in $\gamma_t{}^{EB}$ will increase it. Consider the effect of varying $\gamma_\sigma{}^B$, since it contributes the dominant covalent term in Table XV. If we assume Eq. (24) to be rigorously appropriate to the experimental f_σ value, we can obtain an experimental $\gamma_\sigma{}^B$ which is four times the value appearing in Tables IX and XI. Inserting this into Eq. (38) yields a e_σ covalent term of ~-1100 cm^{-1} which, holding all other contributions constant, leads to a 10 Dq of only $\sim$1000 cm^{-1}. Thus, a mismatch between a γ value and a set of matrix elements can severely affect a computed 10 Dq. The behavior of these matrix elements determines the result; they could be of sufficient accuracy to yield qualitatively correct γ values while being quite unsatisfactory for an estimate of 10 Dq.

As stated, Eq. (38) should be evaluated for the exact self-consistent Hamiltonian. In the absence of such computations, we can again gain some indication of the sensitivity of the results to matrix element behavior by evaluating Eq. (38) for 10 Dq with the bonding $\langle\varphi|H|\chi\rangle^B$ matrix elements of Eq. (48). Replacing a $\langle\varphi|h_0|\chi\rangle^B$ integral by $\langle\varphi|H|\chi\rangle^B$ affects both covalent and overlap contributions to 10 Dq. The results of Table X indicate that $\langle\varphi|H|\chi\rangle^B$ is smaller in magnitude than its partner $\langle\varphi|h_0|\chi\rangle^B$ for small covalent mixing (eventually reversing sign for greater covalency), and imply, in turn, a decrease in 10 Dq. Using $\langle\varphi|H|\chi\rangle^B$ values appropriate to the computed γ's, decreases the covalent and overlap contributions by $\sim$2000 cm^{-1} and yields a 10 Dq value $\approx$800 cm^{-1}, again increasing the numerical disagreement with experiment.

As was the case for the f_i's, the present calculations yield a 10 Dq which is appreciably less than experiment. These underestimates may be related; if so, more than a simple underestimate of γ values has occurred. We are involved with the more subtle, and more difficult, matter of matrix element behavior. Unfortunately, we have little evidence of whether a treatment involving the exact one-electron Hamiltonian, will or will not improve the numerical agreement with experiment. It is important to note that approximations were made[10] in estimating the V_L contributions to the $\langle\varphi|h_0|\chi\rangle$ and $\langle\chi|h_0|\chi\rangle$ matrix elements, but it is presently not obvious what repercussions, if any, these had on the results. This will be investigated in the future. Of greatest interest, of course, is the question of whether a treatment, such as the one outlined above, with the exact h, yields reasonable agreement with experiment. It need not.

IX. OTHER PARAMETERS AFFECTED BY COVALENCY

Other parameters are affected by covalency, including the orbital reduction factors appropriate to the interelectronic Slater F^k (or Racah B and C) integrals of multiplet theory[36] and to g shifts and spin-orbit coupling,[37] and the determination of superexchange[12] interactions and neutron form factors.[38]

In the antibonding approach to the expectation value of a one-electron operator, Op, one compares $\langle\Psi_i{}^A|\mathrm{Op}|\Psi_j{}^A\rangle$ with the free ion $\langle\varphi_i|\mathrm{Op}|\varphi_j\rangle$ value. Quite often, a reduction factor defined by

$$k=\langle\Psi_i{}^A|\mathrm{Op}|\Psi_j{}^A\rangle/\langle\varphi_i|\mathrm{Op}|\varphi_j\rangle \tag{57}$$

is introduced to account for observed differences. Such a factor is, in principle, straightforward to obtain but is, in practice, almost never evaluated because of difficulties associated with the $\langle\chi|\mathrm{Op}|\varphi\rangle$ and $\langle\chi|\mathrm{Op}|\chi\rangle$ matrix elements. The one case where Eq. (57) has been evaluated occurs for Stevens' orbital reduction factor, k,[37,14] appropriate to

$$\langle\Psi_i|l_m|\Psi_j\rangle=k_{i,j,m}\langle\varphi_{3di}|l_m|\varphi_{3dj}\rangle, \tag{58}$$

[36] See, e.g., J. C. Slater, *Quantum Theory of Atomic Structure* (McGraw-Hill Book Company, Inc., New York, 1960), Vol. I.
[37] K. W. H. Stevens, Proc. Roy. Soc. (London) A219, 542 (1953).
[38] J. Hubbard and W. Marshall (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where l_m is some component of the orbital angular momentum **l**. The $\varphi\chi$ and $\chi\chi$ contributions to k can be included and k tends to be less than one.

For other operators it is often assumed that the reduction effects are adequately predicted by the approximate relation

$$\langle \Psi_i^A | \mathrm{Op} | \Psi_j^A \rangle = N_i N_j \langle \varphi_{3di} | \mathrm{Op} | \varphi_{3dj} \rangle , \qquad (59)$$

where the N's are the normalization constants of Eq. (2). The $\langle \varphi | \mathrm{Op} | \chi \rangle$ and $\langle \chi | \mathrm{Op} | \chi \rangle$ terms are neglected, because they are inconvenient to evaluate and are, hopefully, negligible. Detailed inspection of the Stevens' orbital reduction factor suggests that this approximation is inadequate. Nevertheless, necessity led Tinkham to rely on such a normalization estimate when discussing spin-orbit effects.[14] Sugano and Shulman also used this approach when dealing[10] with the experimentally observed[11] reduction of the Racah B and C integrals from their free ion values (here, the normalization enters in the fourth power because two-electron terms are involved).

In the (multielectron) unpaired bonding approach, the simple one-electron picture of the traditional model is lost. As discussed earlier, the overlap effects arise from the bonding-antibonding pairs and covalent contributions come from the unpaired bonding electrons. Now, if one ignores the $[\varphi^*(\mathbf{r})\chi(\mathbf{r})]$ overlap density terms, overlap effects increase the charge and spin on the Ni^{2+} ion. Covalent mixing also increases the charge but decreases the spin (i.e., increases minority spin). Our picture of reduction effects therefore depends on whether we are dealing with a charge or a spin-dependent effect.

The complications introduced by the multielectron considerations increase the size of the computations, but introduce no formal difficulties. When dealing with a diagonal one-electron operator matrix element one need only sum over paired bonding-antibonding, and unpaired bonding contributions. The neutron magnetic form factor is an interesting example of this case. For an off-diagonal matrix element [e.g., between the ground and excited states of Eq. (25)] we must remember that as one changes the occupation of antibonding orbitals, the bonding orbitals are also changed by the change in covalency [it was this which invalidated Eq. (26) as an expression for $10\,Dq$]. Such variations cause the orbitals of one state to lose their orthonormality with respect to orbitals of the other. These complications are easily managed computationally by using a formalism such as Löwdin's[39] for dealing with nonorthogonal functions or by evaluating the matrix element in terms of the antibonding holes (s). The difficulties associated with $\langle \varphi | \mathrm{Op} | \chi \rangle$ and $\langle \chi | \mathrm{Op} | \chi \rangle$ matrix elements, mentioned above, also occur in the present approach. Detailed estimates of orbital reduction effects will not be attempted in this paper. Let us briefly consider an example using the incorrect normalization correction to make an estimate of a reduction effect. In Sugano and Shulman's calculations, the competition between the overlap and the (dominant) antibonding covalent contributions, led to reductions in the Racah B and C integrals equal to those actually observed. If one were to consider such simple normalization corrections for the unpaired bonding results, one finds that the antibonding overlap terms lead to a 3 and 4% *enhancement* of the Racah B and C integrals and a negligible covalent contribution from the unpaired bonding electrons.

Similarly, the small magnitude of the unpaired bonding covalent mixing leads to poor agreement [using, say, Eq. (59)] with the various experimental "k" values.

On several occasions in this section, we have cited difficulties associated with evaluating $\langle \chi | \mathrm{Op} | \varphi \rangle$ and $\langle \chi | \mathrm{Op} | \chi \rangle$ integrals. As this matter has implications for any more exact theory of crystal field effects, let us consider it briefly here. One is frequently interested in the highly singular r^{-3} operator appropriate to fine structure and hyperfine interactions. Matrix elements of this, and other, operators often nearly diverge because the LCAO wave functions were not required to have the proper singular behavior in the vicinity of each of the nuclei present (cf. the previous discussion of transferred hyperfine effects). Thus, even if the $\langle \varphi | \mathrm{Op} | \chi \rangle$ and $\langle \chi | \mathrm{Op} | \chi \rangle$ integrals are well behaved, their computed numerical values could be quite unrealistic. To our knowledge this matter has not been investigated and correcting for it promises to be a most interesting problem. Until this is done, one cannot claim complete quantitative understanding of the various orbital reduction effects.

X. CONCLUDING DISCUSSION

A large number of observations, some major and some minor, have been made, concerning the application of the LCAO molecular orbital method to crystal field theory, in this paper. Perhaps the most important of these is the fact that it is the *unpaired bonding* electrons and not the antibonding electrons, which make covalent contributions to physical observables. The importance of this observation is largely due to the *intrinsic* (hence, quantitative) difference in the nature of the two types of covalent mixing (based on diagonalizing the one-electron Hamiltonian). In addition, we saw indications that a theory based on the incomplete Hamiltonian h_0, is internally inconsistent, hence inadequate. The results, obtained with h_0, were also seen to be in poor agreement with experiment, the covalent contributions being significant but by no means dominant. We are not optimistic that the same theory, evaluated with the exact H-F Hamiltonian, will recapture good agreement with all experimental parameters. (These observations are, of course, made for the molecular orbital theory. The alternative Heitler-London approach to the crystal field prob-

[39] P. O. Löwdin, Phys. Rev. **97**, 1474, 1490, 1509 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

lem[13] is being studied by Rimmer and Hubbard.[40] The Heitler-London method introduces covalency through configuration mixing and thereby loses the simplicity of the single determinant description of the MO theory.)

The fact that the unpaired bonding orbitals make the only covalent contributions to an observable, is not surprising. It has been recognized for some time that, when dealing with Hartree-Fock single determinants, any mixing between a pair of *occupied* orbitals in no way affects the expectation value of an operator.[36] This immediately implies that the covalency of the simultaneously occupied orbitals of a bonding-antibonding pair, cannot in any way, affect the prediction of any observable, for the (NiF_6) cluster states considered here.

Our understanding of orbital reduction effects and of the covalent "delocalization" of charge and spin differs in the two approaches. In particular, the covalency of a bonding-antibonding pair conserves charge on the metal ion while that of the unpaired bonding orbital involves a shift of charge from the ligands to the metal. Such differences in the *nature* of the two types of covalent mixing, have quantitative repercussions on the estimates of γ values (cf. Table XII). This is the most immediate quantitative reason for our preoccupation with unpaired bonding effects in this paper. As noted, the differences between computed paired and unpaired covalent mixing do not destroy the required spin symmetry of the cluster (i.e., it is an eigenfunction of S^2). The unpaired bonding approach may be viewed as the proper application of *restricted* H-F theory to the current cluster model for $KNiF_3$. Spin symmetry is maintained and what is more, if we apply the variation principle to the total energy of the cluster (with respect to the covalent mixing parameters), subject to the (RHF) requirement that the two types of covalent mixing be identical, one of necessity obtains the unpaired bonding (or antibonding hole) values. (The discussion for Cr^{3+} has, of course, encompassed UHF theory and the breakdown of spin symmetry.)

We also noted that a theory based on the mixing of free ion metal φ and ligand ψ orbitals and based on the approximate Hamiltonian, h_0, appeared unsatisfactory (quite aside from the fact that quantitative agreement with experiment was not obtained) for we saw serious internal inconsistencies in the covalent mixing estimated for bonding-antibonding pairs. The effects did not seem so severe for unpaired bonding γ values, an observation which is probably misleading.

[40] D. Rimmer and J. Hubbard (unpublished), who use the configuration interaction method of Keffer *et al.* (Ref. 13).

Having failed to produce even fair agreement with experiment, the question arises whether the same theory, evaluated for the exact self-consistent Hamiltonian h, will or will not produce agreement with experiment. We do not expect the current (free ion ψ and φ) cluster model to yield good agreement (say to 20%) for all the parameters considered in this paper. Should such a failure prove to be the case, there will again arise the old question of the variation of atomic orbital character away from the free ion behavior assumed in the present LCAO approach. Such a variation can take several forms: first, there is the matter of bulk expansion or contraction of the metal[17,41] ion (or ligands); secondly, there is the question of the extent to which $3d\varphi_t$ and φ_3 orbitals differ radially[42]; thirdly, there can be variation in orbital behavior from one multielectron state to the other; fourthly, one can allow different bonding and antibonding ψ's to have different radial parts for their φ and χ basis orbitals (involving us almost inevitably with a UHF theory); lastly, we have to consider the variation of orbitals centered on one nucleus when in the immediate vicinity of another, a matter important to orbital reduction and hyperfine effects. A bulk expansion of the metal φ orbitals may be expected with the covalent increase of charge occuring at the Ni site. This would improve numerical agreement with some experiments but it remains to be seen if this effect is fundamental to the current disagreement with experiment. Other questions which arise, include the role of correlation effects, one's ability to adequately define V_{ext} and the justification of the cluster model itself.

Finally, it is imperative that a proper crystal field cluster treatment, involving the exact Hamiltonian, be done. In this way, a number of the questions of the present paper would be resolved. We believe a calculation of this sort to currently be attainable with reasonable accuracy.

Note added in proof. In a recent publication, E. Šimánek and Z. Šroubek [Physica Status Solidi 4, 251 (1964)] have considered unpaired bonding effects for the same system.

ACKNOWLEDGMENTS

We are pleased to acknowledge conversations with S. Geschwind, D. Rimmer, L. R. Walker, and Y. Yafet.

[41] L. Orgel, J. Chem. Phys. **23**, 1824 (1955).

[42] Papers endeavoring to estimate this effect include R. E. Watson, Phys. Rev. **117**, 724 (1960); and R. E. Watson and A. J. Freeman, *ibid.* **120**, 1134 (1960). (There is an error associated with the cubic field term used in the first paper causing it to be approximately twice as strong as intended.)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Magnetic Moments and Unpaired Spin Densities in the Fe-Rh Alloys*

G. SHIRANE AND R. NATHANS
Brookhaven National Laboratory, Upton, New York

AND

C. W. CHEN
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania

(Received 22 January 1964)

The distribution of magnetic moments in the Fe-Rh system has been investigated by the neutron diffraction technique in the composition range between 35 and 50 at.%Rh. These alloys have chemical order of CsCl type; the body-corner positions are occupied by FeI atoms and the body centers by Rh and FeII atoms. The magnetic moments in ferromagnetic alloys containing 35, 40, and 48% Rh are $\mu_{FeI}=3.1\ \mu_B$, $\mu_{FeII}=2.5\ \mu_B$, and $\mu_{Rh}=1.0\ \mu_B$ at 25°C. The moment of the FeI atom in the 50% Rh alloy, which is antiferromagnetic at room temperature, increases to $3.3\ \mu_B$. A detailed study of the unpaired spin density distribution in the 48% Rh alloy was made by the polarized beam technique. While the FeI atom has a nearly spherical density distribution, that of the Rh atom shows a strong tendency towards e_g symmetry. The coherent nuclear scattering amplitude of Rh was determined to be $0.585\pm0.005\times10^{-12}$ cm.

I. INTRODUCTION

IN a previous paper[1] we have reported a systematic investigation of the Fe-Rh alloys by the Mössbauer technique. Up to 52% Rh these alloys possess a body-centered cubic lattice in which there is chemical order of CsCl type when the Rh concentration exceeds 20%. In the ordered alloys, iron atoms located at the body corners are labeled FeI and those in the Rh sites at the body centers FeII (see the insert to Fig. 1). The system exhibits two interesting magnetic properties. First, fairly large moments are observed over a wide concentration range,[1,2] as shown in Fig. 1. Second, a ferromagnetic-antiferromagnetic transition[3,4] occurs in alloys containing around 50% Rh. The transition temperature increases with increasing Rh concentration, and is observed to be 65°C in the 50% Rh alloy. The room temperature phase changes from ferromagnetic to antiferromagnetic between 49% and 50% Rh.

The Mössbauer measurements[1] revealed two distinct hyperfine fields in ordered alloys, corresponding to FeI and FeII. For example, H_i(I) and H_i(II) were found to be 277 and 384 kOe, respectively, in an alloy containing 48% Rh. It was suggested that FeII, which is surrounded by 8 Fe nearest neighbors, may have a smaller magnetic moment than FeI despite its higher hyperfine field. The main interest in this alloy system therefore lies in the composition dependence of both the moments and the hyperfine fields of FeI, FeII, and Rh.

A neutron diffraction study of the FeRh alloy was first carried out by Bertaut, de Bergevin, and Roult.[5] Using powder data, they reported $\mu_{Fe}=2.8\ \mu_B$ and $\mu_{Rh}=0.8\ \mu_B$ for the 53% Rh alloy in the ferromagnetic region above 60°C. The present work describes neutron diffraction measurements of the magnetic moments of FeI, FeII, and Rh in ordered alloys, and a detailed study of the spin density distribution of Fe and Rh in the 48% Rh alloy. A study of the 50% Rh alloy, which is antiferromagnetic at room temperature, was also made.

II. EXPERIMENTAL

A. Samples

The alloys used in this work are those previously investigated by the Mössbauer technique. All alloys

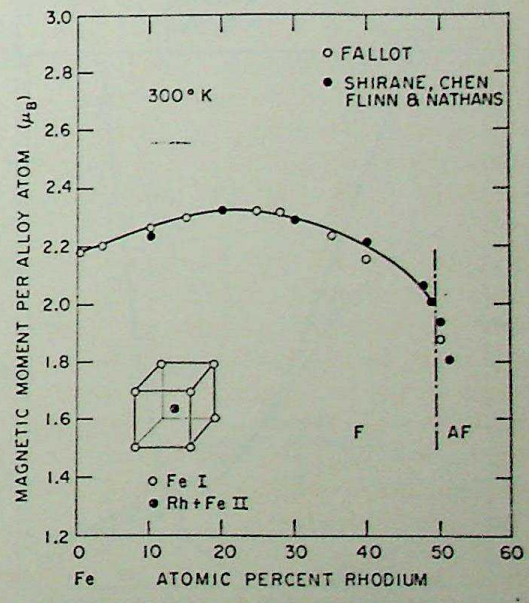


FIG. 1. Average atomic moments of Fe-Rh alloys at 300°K. The values in the antiferromagnetic phase were obtained by extrapolation of the data obtained for the high-temperature phase.

* Work supported by the auspices of the U. S. Atomic Energy Commission and the Advance Research Project Agency.

[1] G. Shirane, C. W. Chen, P. A. Flinn, and R. Nathans, Phys. Rev. 131, 183 (1963).

[2] M. Fallot, Ann. Phys. (Paris) 10, 291 (1938); M. Fallot and R. Hocart, Rev. Sci. 77, 498 (1939).

[3] F. de Bergevin and L. Muldawer, Compt. Rend. 252, 1347 (1961).

[4] J. S. Kouvel and C. C. Hartelius, J. Appl. Phys. Suppl. 33, 1343 (1962).

[5] F. Bertaut, F. de Bergevin, and G. Roult, Compt. Rend. 256, 1668 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Physical properties of Fe-Rh alloys. The moment value of 50% Rh (*) is obtained by extrapolation from the ferromagnetic phase above 65°C.

Composition	48% Rh	50% Rh
Lattice constant	2.989 Å	2.986 Å
Curie temperature	500°C	405°C
Magnetic moment per atom, 300°K	2.07 μ_B	1.94 μ_B*
Néel temperature	...	65°C

were prepared from high-purity metals by levitation melting with subsequent heat treatments as described in the previous paper.[1] The chemical analysis agreed with the nominal composition within 0.5 at.%. Single crystals large enough for neutron diffraction study were obtained by levitation melting of the 48 and 50% Rh alloys. Efforts to grow single crystals of the 25% Rh alloy by the electron beam technique were not successful. This may be due to the existence of a γ phase at high temperatures. An attempt to obtain single crystals of the 35% Rh alloy by the stress annealing technique was also unsuccessful.

Thus polycrystalline samples containing 35 and 40% Rh were used to study the moments of FeI and FeII. Detailed studies of the ferromagnetic and antiferromagnetic phases were carried out on the 48% Rh and 50% Rh single crystals, respectively. The magnetic properties of these two alloys are listed in Table I.

Interpretation of the magnetic diffraction data requires a precise knowledge of the chemical order in these alloys. Previous Mössbauer measurements, coupled with x-ray powder patterns, had revealed pronounced order in all the alloys containing more than 20% Rh. A careful study of the 50% Rh alloy, to be described later, showed that ordering is complete. Somewhat less accurate measurements on the 35% and 48% alloys gave a similar result.

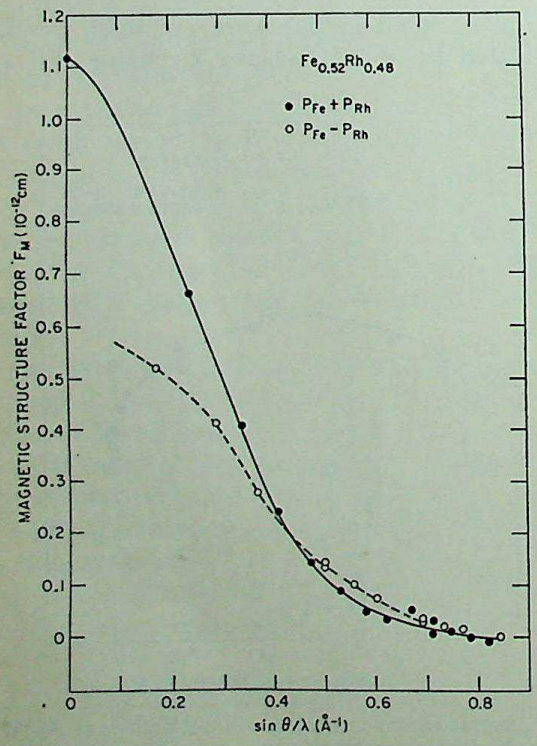


FIG. 2. The observed magnetic structure factors F_M, of $Fe_{0.52}Rh_{0.48}$. Reflections with $h+k+l=2n$ are shown by solid circles, those with $h+k+l=2n+1$ by open circles. The point at $\sin\theta/\lambda=0$ corresponds to the observed bulk moment of 4.14 μ_B.

B. Polarized Neutron Technique

Except for the study of the antiferromagnetic phase containing 50% Rh, all the measurements were made with the use of the polarized neutron technique, which has already been described in detail.[6–9] By reversing the incident neutron polarization, we measure the polarization ratio

$$I_{(+)}/I_{(-)}=(1+\gamma)^2/(1-\gamma)^2,$$

where γ is the ratio of the magnetic to the nuclear structure factor (F_M/F_N). The observed ratios were corrected for incomplete polarization of the incident beam (0.98) and for incomplete polarization reversal (0.98).

It is also necessary to make allowance for multiple scattering processes within the crystal. It has been shown[9] that multiple scattering can introduce serious errors in the measured polarization ratios. In order to detect and eliminate this effect, the crystals were mounted on a goniometer in a magnetizing field of 10 kOe, and the polarization ratio was studied as a function of the azimuthal angle around the scattering vector.

III. SPIN DENSITY IN FERROMAGNETIC $Fe_{0.52}Rh_{0.48}$

Single crystals were cut into the shape of pillars, roughly 2×3 mm in cross section and 15 mm in length, with their long axes along [110] and [100], respectively. In addition, a smaller cylindrical [110] crystal (1-mm diam), which was cold worked, was used to check for possible extinction effects in strong reflections. Extinction effects were also investigated by mis-setting the Bragg angle.[6] Even for undistorted crystals, extinction corrections were proved to be negligible except for the two strongest reflections.

Intensity measurements were made out to an angle corresponding to $\sin\theta/\lambda=0.85$ Å^{-1}, with a wavelength of 1.05 Å. The number of reflections totaled 26, and all except two, (321) and (421), are contained in the two zones. The results are shown in Fig. 2 and Table II. In calculating the magnetic structure factor F_M from the relationship $\gamma=F_M/F_N$ we used $b_{Fe}=0.951$[9] and $b_{Rh}=0.585\times10^{-12}$ cm, the latter value being determined from the 50% Rh alloy (see Sec. V). The errors quoted

[6] R. Nathans, C. G. Shull, G. Shirane, and A. Andresen, Phys. Chem. Solids **10**, 138 (1959).
[7] R. Nathans, Suppl. J. Appl. Phys. **31**, 350 (1960).
[8] S. J. Pickart and R. Nathans, Phys. Rev. **123**, 1163 (1961).
[9] C. G. Shull and Y. Yamada, J. Phys. Soc. Japan **17**, Suppl. B III, 1 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in Table II are the standard deviations of repeated measurements.

In the CsCl type of structure, the structure factors F_M are divided into two kinds,

$$F_M = p_{Fe} + p_{Rh} \quad \text{for} \quad h+k+l=2n,$$
$$F_M = p_{Fe} - p_{Rh} \quad \text{for} \quad h+k+l=2n+1,$$

where p_{Fe} and p_{Rh} are the magnetic scattering amplitudes of Fe and Rh, respectively, and are equal to the product of the magnetic moment μ and the form factor f,

$$p_{Fe} = \mu_{Fe} f_{Fe}, \quad p_{Rh} = \mu_{Rh} f_{Rh}.$$

Figure 2 clearly shows that these two groups of reflections lie on two separate curves which cross each other around $\sin\theta/\lambda = 0.45$ Å^{-1}. This implies that the rhodium form factor f_{Rh} becomes negative at this point, in agreement with recent calculations by Watson and Freeman.[10] Thus it falls off much more rapidly as a function of the scattering angle than the form factor of Fe, which becomes negative only above $\sin\theta/\lambda = 0.80$ Å^{-1}.

TABLE II. Magnetic structure factors F_M, in units of 10^{-12} cm, per CsCl type unit cell of $Fe_{0.52}Rh_{0.48}$.

hkl	$F_M(\pm 0.001)$	hkl	$F_M(\pm 0.002)$
100	0.516	011	0.658
111	0.410	200	0.400
120	0.275	211	0.237
300	0.136	022	0.141
122	0.144	130	0.090
311	0.104	222	0.053
230	0.076	321	0.035
322	0.039	400	0.052
140	0.037	033	0.009
133	0.023	411	0.030
421	0.016	240	0.012
500	0.003	233	0.000
430	0.004	422	−0.010

These data were analyzed in two ways. The first is the direct comparison of observed and calculated structure factors, and the second involves the Fourier inversion of the data. We shall describe the second approach first, since it gives a more visible picture. This process of Fourier inversion was used by Pickart and Nathans[8] for Fe_3Al and Shull and Yamada[9] for Fe. Our preliminary analysis was made with a Fourier projection on the (110) plane, but a heavy overlap due to an extended wave function of Rh made the three-dimensional Fourier analysis desirable.

Following conventional crystallographic notation, we have

$$\rho(xyz) = (1/V) \sum F_{hkl} \cos 2\pi(hx+ky+lz),$$

where ρ is the magnetic spin density and V is the unit cell volume. The summation has been carried out on the computer. Figure 3 illustrates a portion of the (110) plane intersecting the two atoms in the unit cell. This Fourier map clearly illustrates the lobe along the [001] axis for Rh, the characteristic of e_g symmetry. On the other hand, the Fe density distribution corresponds more closely to spherical symmetry. Figure 4 illustrates the moment density along two principal directions, [001] and [111], the latter being directed towards the nearest neighbors. This figure shows that the Fe spin density also shows some deviation towards e_g symmetry, in that the density along [001] is slightly higher than that along [111] at the equivalent distance.

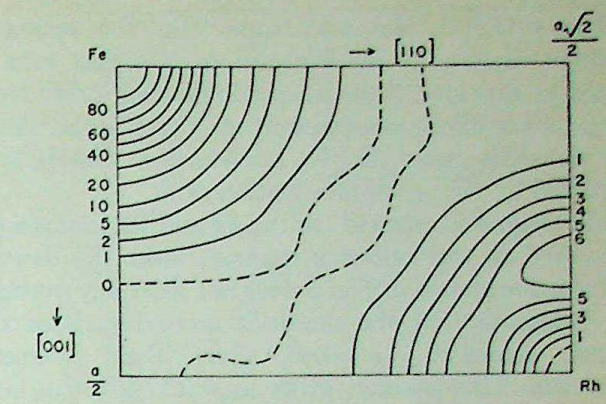


FIG. 3. Magnetic moment density distribution in the (110) plane. The scale is taken to give 100 at the Fe nucleus, which corresponds to 6.82 μ_B/Å^3.

It is emphasized that the large difference in the angular dependence of the form factors, f_{Fe} and f_{Rh}, has a significant effect on their Fourier contours. Available data within the given angle of $\sin\theta/\lambda = 0.85$ Å^{-1} allow an excellent resolution of the Rh spin density distribution to be made since its form factor crosses zero at

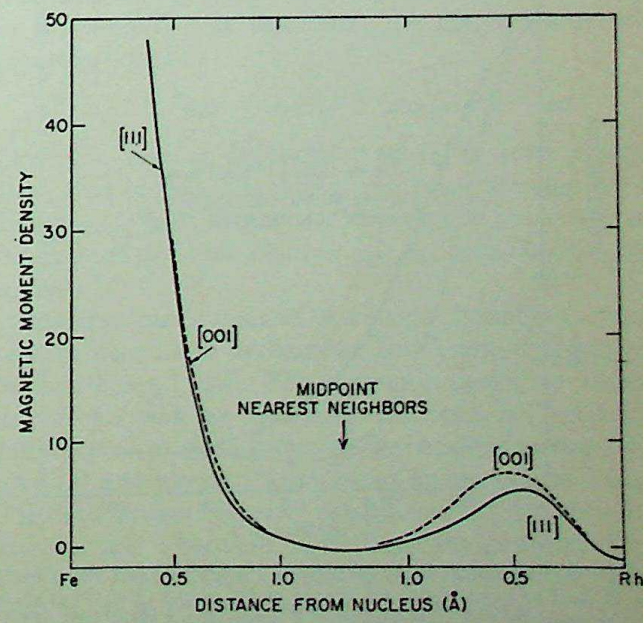


FIG. 4. Magnetic moment density distribution in the [001] and [111] directions as a function of distance from the nuclei. The scale 10 corresponds to 0.68 μ_B/Å^3. [111] corresponds to the nearest-neighbor direction.

[10] R. E. Watson and A. J. Freeman (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$\sin\theta/\lambda = 0.45$ Å^{-1}. We may note that the magnetic moment density at the Rh nucleus is almost zero as would be expected if the magnetic moment arises from the presence of unpaired d electrons. On the other hand, the densities near the Fe nucleus are considerably distorted by the series termination error.

The magnetic moments of the Fe and Rh atoms were estimated in the following manner. Since the density distribution shown in Fig. 3 does not show any overlap, the integration of the densities around each of the atoms, namely $\sum \rho(r) 4\pi r^2$, gives their magnetic moments. This method gives $\mu_{Fe} = 3.17 \pm 0.10\ \mu_B$ and $\mu_{Rh} = 0.97 \pm 0.10\ \mu_B$. These values are subject to errors which result from the fact that only a finite number of reflections are used in the construction of a Fourier map. The effect of this series termination error on magnetic density was studied in detail by Pickart.[11] From this study we can estimate that in the present case the error is not more than a few percent for Fe. Another estimate was made directly from the data shown in Fig. 2. We draw a smooth curve through the observed points, taking into account the asphericity. Then for a given observed value of $p_{Fe} + p_{Rh}$, that of $p_{Fe} - p_{Rh}$ can be estimated with reasonable accuracy. In this way we can obtain p_{Fe} and p_{Rh} uniquely as a function of $\sin\theta/\lambda$. Then $p_{Fe} = \mu_{Fe} f_{Fe}$ were compared with the form factor of Fe metal.

If we now assign a value of 3.14 μ_B to the FeI moment, it may be seen from Fig. 5 that the form factor thus obtained is almost identical to that of Fe metal as

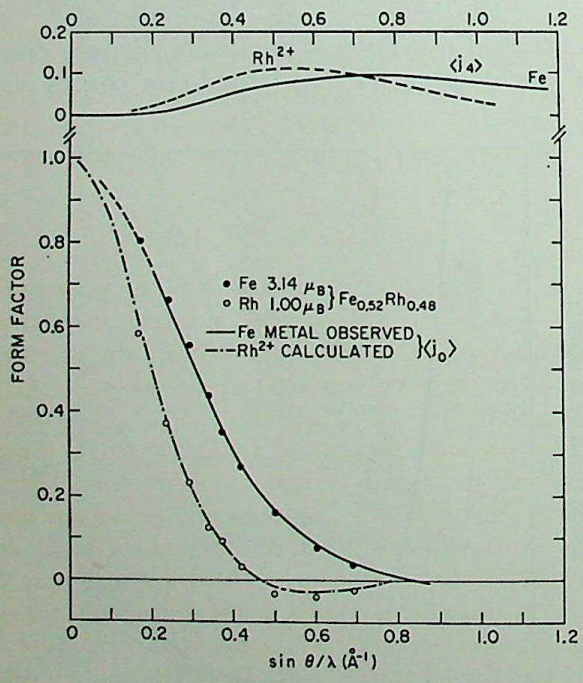


FIG. 5. Estimated form factors of Fe and Rh, as described in the text. The observed form factor of Fe metal (Ref. 9) is shown by the solid line and the calculated form factor (Ref. 10) of Rh^{2+} by the broken line. The calculated aspherical terms $\langle j_4 \rangle$ are shown at the top.

[11] S. J. Pickart, Acta Cryst. 16, 174 (1963).

determined by Shull and Yamada[9] over a quite wide angular range. This Fe moment, combined with the observed bulk moment of 4.14 μ_B, gives $\mu_{Fe} = 3.14\ \mu_B$, $\mu_{Rh} = 1.0\ \mu_B$.

At present, theoretical form factors[10] for Rh are available only for Rh^{2+} and Rh^{3+}. The observed form factor f_{Rh} turns out to be in excellent agreement with the calculated value for Rh^{2+}. The almost perfect agreement shown in Fig. 5 is probably fortuitous. In the above process, we have neglected the fact that a small fraction (4%) of the Rh sites are occupied by FeII atoms. If we assume the moment of FeII atoms is 2.5 μ_B, as will be shown in the next section, the true rhodium moment in $Fe_{0.52}Rh_{0.48}$ should be 0.94 μ_B, and the form factor should be slightly modified.

The aspherical moment distribution, particularly pronounced in Rh, was already apparent in the Fourier

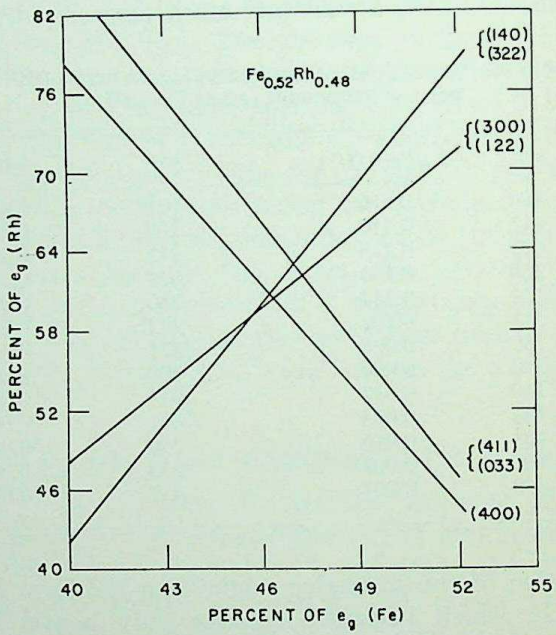


FIG. 6. Graphical determination of the fractions of e_g character for Fe and Rh in $Fe_{0.52}Rh_{0.48}$. 40% e_g corresponds to spherical symmetry.

densities shown in Figs. 3 and 4. A qualitative estimate was made by comparison of observed and calculated structure factors of three pairs of reflections, shown in Fig. 6, and (400) which shows the largest deviation from the spherical part. The form factor for e_g symmetry is given by

$$f(e_g) = \langle j_0 \rangle + \tfrac{3}{2} A \langle j_4 \rangle ,$$

$$A = \frac{h^4 + k^4 + l^4 - 3(h^2k^2 + k^2l^2 + l^2h^2)}{(h^2 + k^2 + l^2)^2} .$$

The calculated values[10,12] for $\langle j_4 \rangle$ are shown in the top part of Fig. 5. By a graphical method, the best solution for the percentage of e_g character in the spin density

[12] R. E. Watson and A. J. Freeman, Acta Cryst. 14, 27 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

distribution was found to be $e_g(\mathrm{Fe})=47\pm1\%$, $e_g(\mathrm{Rh})=62\pm3\%$.

Spherical symmetry corresponds to 40% e_g and 60% t_{2g} character.

IV. MOMENT DISTRIBUTION IN ALLOYS

Having established the values of the magnetic moments in the 48% Rh alloy, it is now of great interest to study the composition dependence of these moments, and in particular, that of FeII, namely Fe atoms on the Rh sites. These features were studied with polycrystalline samples of the 35 and 40% Rh alloys consisting of rectangular pillars 2×5×18 mm. Both the polarized and unpolarized beam techniques were employed in these experiments. In the case of the former, a correction for depolarization of the neutron beam within the polycrystalline samples is

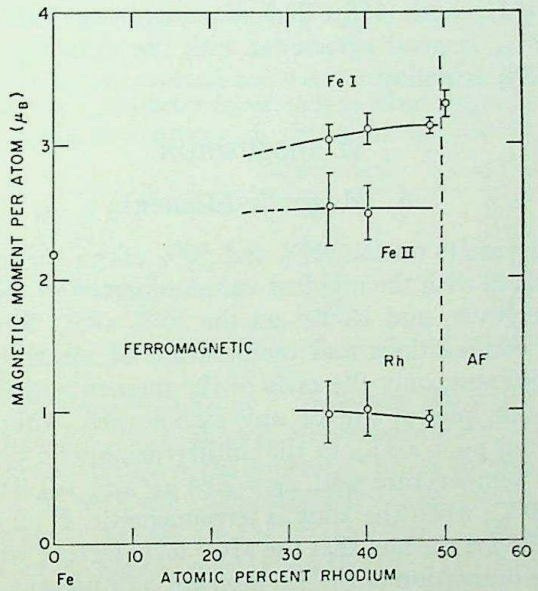


FIG. 7. Moment distributions in Fe-Rh alloys at room temperature.

required. With these in an applied field of 11 kOe, this correction turned out to be relatively small. The degree of chemical order in the 35% Rh sample was checked, and the order parameter was found to be at least 0.95.

The data were collected from a few reflections, (110), (200), (211), and (100). From these we can determine $F_M(+)=p_{\mathrm{FeI}}+p_{\mathrm{FeII}}+p_{\mathrm{Rh}}$ and $F_M(-)=p_{\mathrm{FeI}}-(p_{\mathrm{FeII}}+p_{\mathrm{Rh}})$. These, combined with the observed total moment by magnetic measurements, give the individual moments uniquely if we assume that the form factors of Fe and Rh determined for the 48% Rh alloy are unchanged in the other alloys.

The results are shown in Fig. 7. Large uncertainties, particularly in the moment of FeII, were introduced because of poor statistics in the powder diffraction data. It is quite clear, however, that μ_{FeII} is definitely smaller than μ_{FeI} despite its higher hyperfine field.

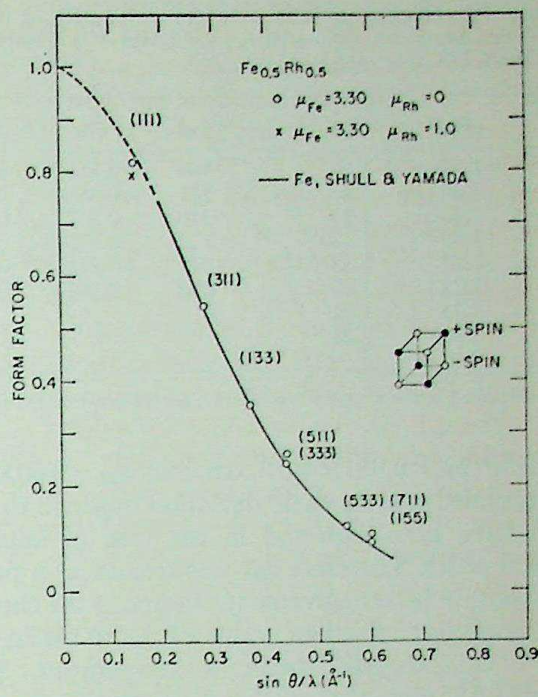


FIG. 8. Form factor of Fe in the antiferromagnetic phase of $Fe_{0.50}Rh_{0.50}$ at room temperature. The assumption that $\mu_{Rh}=1.0\,\mu_B$ does not introduce any noticeable change in the Fe form factors except for the (111) reflection.

V. ANTIFERROMAGNETIC PHASE IN $Fe_{0.5}Rh_{0.5}$

The basic magnetic structure of this phase was reported by Bertaut *et al.*[5] and is shown as an insert to Fig. 8. This is the so-called *G*-type structure,[13] in which every positive Fe spin is surrounded by six negative Fe spins. The magnetic unit cell edge is thus double that of the chemical cell. The Rh atom, which is located at the body-centered position, is now surrounded by four positive and four negative Fe spins. Thus the Rh moments may not be ordered. Even if they are ordered in the same *G*-type structure, the magnetic intensity is proportional to

$$F_M^2=(\mu_{\mathrm{Fe}}f_{\mathrm{Fe}})^2+(\mu_{\mathrm{Rh}}f_{\mathrm{Rh}})^2.$$

Since the angular dependence of f_{Fe} is quite different from that of f_{Rh}, it is possible in principle to solve for both μ_{Fe} and μ_{Rh}. In practice, however, F_M^2 is extremely insensitive to small changes of μ_{Rh} when $\mu_{\mathrm{Rh}}f_{\mathrm{Rh}}<0.1\,\mu_{\mathrm{Fe}}f_{\mathrm{Fe}}$.

The specimen consisted of a single crystal in the shape of a cylinder 1 mm in diameter and 5 mm in length with its axis along [110]. The magnetic peaks are indexed as *hkl* all odd by doubling the unit cell. Accurate intensity data were collected for both magnetic (*h*,*k*,*l* all odd) and nuclear peaks using an unpolarized beam. The absorption of the sample was relatively high ($\mu=4.2$ cm^{-1}) because of its large Rh content. A preliminary comparison of the observed structure factors with those calculated from the expressions

$$F_{\mathrm{cal}}=(b_{\mathrm{Fe}}\pm b_{\mathrm{Rh}})e^{-B(\sin\theta/\lambda)^2},$$

[13] E. O. Wollan and W. C. Koehler, Phys. Rev. 100, 545 (1955).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE III. Nuclear structure factors of $Fe_{0.5}Rh_{0.5}$, in units of 10^{-12} cm. Statistical errors in F_{ob} are about 1%. Temperature factor $B=0.433$ Å^2 is included in F_{cal}.

hkl	F_{ob}	F_{cal}	hkl	F_{ob}	F_{cal}
011	1.47[a]	1.50	100	0.356	0.362
200	1.46	1.46	111	0.358	0.354
211	1.43	1.43	300	0.332	0.329
022	1.39	1.40	122	0.326	0.329
222	1.32	1.33	311	0.324	0.321
400	1.28	1.26	322	0.293	0.298
033	1.25	1.24			
411	1.24	1.24			

[a] Not included in the analysis because of a probable extinction effect.

in which[9] $b_{Fe}=0.951\times10^{-12}$ cm and $b_{Rh}=0.60\times10^{-12}$ cm,[14] revealed a systematic deviation opposite to what would have been expected in the case of imperfect chemical order. Therefore b_{Rh} was treated as a parameter to obtain better agreement. Assuming the chemical order is perfect, the new value of $b_{Rh}=0.585\pm0.005\times10^{-12}$ cm, together with a temperature factor $B=0.433$ Å^2, give the excellent agreement shown in Table III.

If chemical order is not complete, or the Rh content not 50%, it is clear that b_{Rh} would become even smaller. The above value of the temperature factor B was used to correct the observed magnetic structure factors. Magnetic intensities were also corrected for $\lambda/2$ contamination ($\sim$0.2%). Intensity data given in Table IV are average values of equivalent reflections. In view of the nuclear intensity agreement, it was assumed that no extinction correction was needed.

The magnetic reflections were first analyzed on the assumption that $\mu_{Rh}=0$. The Fe moment was determined from the (311) and (133) reflections using the form factor of metallic Fe. This gave

$$\mu_{Fe}=3.30\,\mu_B.$$

The form factor f_{Fe} was then calculated for all the reflections using this moment value. There appears to be a slight systematic deviation of the observed form factor from the solid line representing that of metallic Fe as shown in Fig. 8. A part of this might be due to a systematic error introduced by an inaccurate temperature factor correction. (The polarized beam technique is free from this error since it is always the ratio, F_M/F_N, which is measured.)

Some estimate of the Rh moment can be made on the basis of the innermost magnetic peak. If we were to assume that the Rh moment remains 1.0 μ_B in the antiferromagnetic state, then we obtain the value of the Fe form factor indicated by the cross in Fig. 8, which falls definitely below the extrapolated curve. Hence, our data does suggest that if the aligned moment on the Rh atoms in this magnetic state is not zero, it is at least lowered. However, the determination of the Fe moment is almost unaffected by this uncertainty in the Rh moment.

The asphericity of the Fe density distribution was estimated from two pairs of reflections, (511) and (333), and (711) and (155). The amount of e_g character is 48±3%, in good agreement with the value for FeI in the 48% Rh alloy.

TABLE IV. Magnetic structure factors F_M, in units of 10^{-12} cm, of antiferromagnetic $Fe_{0.5}Rh_{0.5}$ per chemical unit cell. Indices are based on the magnetic unit cell ($a=5.972$ Å). f_{Fe} is calculated on the assumption that $\mu_{Fe}=3.30\,\mu_B$ and $\mu_{Rh}=0$.

hkl	F_M(±0.002)	f_{Fe}
111	0.594	0.817
311	0.396	0.545
133	0.258	0.355
{333	0.178	0.245
{511	0.191	0.263
533	0.089	0.123
{155	0.072	0.099
{711	0.078	0.107

VI. DISCUSSION

A. Magnetic Moments

Our results on the 48% and 50% alloys can now be compared with the moment values reported by Bertaut, de Bergevin, and Roult[5] on the 53% alloy. Since the alloy compositions and temperature of measurements are different, only the ratio of the moments in the two magnetic phases are of any significance. They have reported $\mu_{Fe}=3.3\,\mu_B$ in the antiferromagnetic phase at room temperature and $\mu_{Fe}=2.84\,\mu_B$ and $\mu_{Rh}=0.80\,\mu_B$ at 100°C, when the alloy is ferromagnetic. If allowance is made for the fact that the Mo^{3+} form factor[15] was used in the derivation of the Rh moment, our present result. are seen to be in good agreement with these values The rhodium form factor determined in the present work drops off, as a function of scattering angle, somewhat more slowly than that of Mo^{3+}, as might be expected from their respective positions in the 4*d* series.

It is interesting to compare the magnetic moment of Fe in the antiferromagnetic phase with the Fe moment extrapolated from the ferromagnetic region. At room temperature, the extrapolated total moment is 3.88 μ_B per unit cell (Fig. 1). Assuming that the ratio of μ_{Fe} to μ_{Rh} retains the value found in 48% Rh, we obtain a value of 3.0 μ_B for the Fe moment. The observed moment is 10% higher than this value. This tendency may also be seen in Fig. 7, in which the moments were plotted as a function of composition.

At the transition from the ferromagnetic to the antiferromagnetic phase, the unit cell volume decreases[3] by

[14] C. G. Shull and E. O. Wollan, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1956), Vol. 2, p. 137.

[15] M. K. Wilkinson, E. O. Wollan, H. R. Child, and J. W. Cable, Phys. Rev. **121**, 74 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

1% and the hyperfine field at Fe^{57} decreases by 6%.[1,16] This is another illustration of the fact that, in this system, the hyperfine field is not related in any simple manner to the magnetic moments. We have already shown that FeII in the ferromagnetic phase has a smaller moment, despite its considerably higher hyperfine field, than FeI. We believe this is due to a different degree of polarization of the conduction electrons, with respect to which the hyperfine field is quite sensitive. At present there is, however, no direct experimental evidence to verify this assumption.

B. Form Factors

The form factors of Fe in the rhodium alloys are shown in Fig. 9, together with the previous results given for metallic Fe by Shull and Yamada[9] and for Fe_3Al by Pickart and Nathans.[8] The first striking characteristic is that the spherical parts of the form factors are very similar to each other, although the magnetic moments of the Fe atoms in their various surroundings are quite different, as shown in Table V. However, the asphericity of the density distribution differs considerably.

TABLE V. Magnetic moments and their asphericities in several Fe alloys. 40% e_g character corresponds to spherical symmetry.

	Moment (μ_B)	e_g (%)
Fe Metal	2.18	0.53
Fe_3Al FeI	2.18	0.60
Fe_3Al FeII	1.50	0.48
$Fe_{0.52}Rh_{0.48}$ FeI	3.14	0.47
$Fe_{0.5}Rh_{0.5}$ FeI	3.30	0.48

Let us now compare the spherical part of the observed form factor with the calculated part. The most recent calculations available are the Hartree-Fock calculations with exact exchange polarization treatment by Freeman and Watson[17] for the two configurations $3d^8$ and $3d^64s^2$. Somewhat surprisingly, the form factors for these two configurations turn out to be almost identical.[18] This is in striking contrast to previous calculations by the same authors[12,19] in which they obtained practically identical results for $3d^64s^2$ and $3d^6(Fe^{2+})$, but considerably lower values for $3d^8$.

Shull and Yamada compared their observed form factors of Fe to the latest spin-polarized calculations and obtained a constant "ratio" between the two, namely, that the observed values were 10% higher than the calculated values. They attributed the difference to the polarization of 4s electrons assuming the calculated form factors were correct. Our data for the FeRh alloys as well as the previous results for Fe_3Al deviate from this constant ratio relation at the low-angle reflections with $\sin\theta/\lambda < 0.2$ Å^{-1}. If we include these points for comparison, better over-all agreement is obtained with the earlier calculations of Weiss and Freeman.[20] We believe that, at present, the proper procedure for obtaining information about the conduction electron polarization is to depend only on the experimental intensity observations and their Fourier transform. Our present data for the Fe-Rh alloys are not sufficient to draw significant conclusions concerning this point.

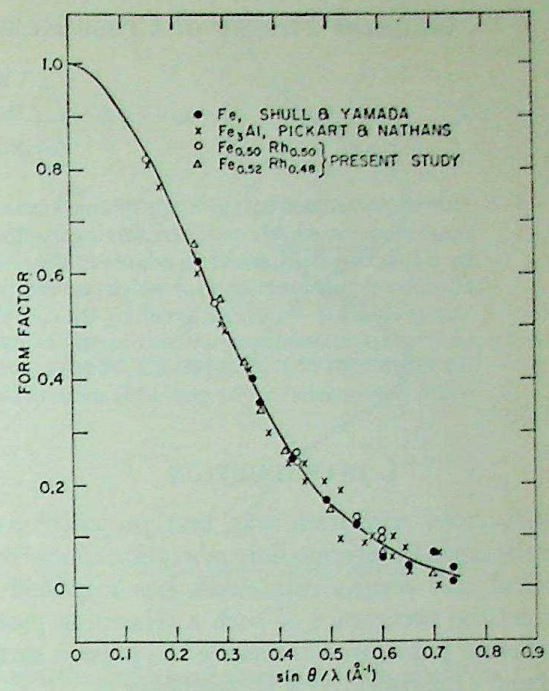


FIG. 9. Iron form factors determined for various alloys.

ACKNOWLEDGMENTS

We are very grateful to C. G. Shull, S. J. Pickart, D. E. Cox, and M. Blume for many stimulating discussions, to R. E. Watson and A. J. Freeman for providing us with their unpublished form factors, and to J. A. Ibers and W. C. Hamilton for the use of their three-dimensional Fourier program.

[16] F. E. Obenshain, L. D. Roberts, and H. H. F. Wegener, Bull. Am. Phys. Soc. 8, 43 (1963).
[17] A. J. Freeman and R. E. Watson (unpublished), quoted by Shull and Yamada (Ref. 9).
[18] If the "true" calculated values of $3d^64s^2$ and $3d^8$ are really so close to each other, we cannot hope to distinguish the $3d$ configuration from the observed form factors.
[19] A. J. Freeman and R. E. Watson, Acta Cryst. 14, 231 (1961).
[20] R. J. Weiss and A. J. Freeman, Phys. Chem. Solids 10, 147 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

General Theory of Cross Relaxation. I. Fundamental Considerations*

W. J. C. GRANT
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received 30 January 1964)

A statistical analysis is made of spin transitions induced by dipole interactions which change the total magnetization while exactly conserving energy. The first-order effect of the dipole operator can be described by a function $\Phi(\omega)$, which is related to the level broadening observed in resonance lines. The second-order effect leads to a function $\chi(\omega)$ which represents the power spectrum of the dipole operator. The cross-relaxation probability $W_{CR}(\omega)$ is given by the convolution of these two functions. W_{CR} is calculated explicitly in various approximations, without appeal to moments. For single-spin flips in magnetically dilute systems, the magnitude of W_{CR} depends linearly on the concentration n. There is a very sharp peak at $\omega=0$ with a width proportional to the geometric mean of the resonance width and of the nearest-neighbor dipole energy.

I. INTRODUCTION

SINCE cross relaxation was first proposed[1] as an explanation of intermediate relaxation times,[2] much theoretical and experimental work has amplified and confirmed the occurrence of such a relaxation process. As a starting point, we summarize the present state of the theory. One considers a Hamiltonian

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_{dip}, \tag{1}$$

where $\mathcal{H}_0$ includes Zeeman and crystal energies, and $\mathcal{H}_{dip}$ is the dipole interaction energy. The cross-relaxation transition probability is given by

$$W(\omega)=(2\pi/\hbar)M^2g(\omega), \tag{2}$$

where M is the transition matrix element arising from $\mathcal{H}_{dip}$, $g(\omega)$ is a probability function peaked at $\omega=0$, and $\hbar\omega$ is the energy that is exchanged between $\mathcal{H}_0$ and $\mathcal{H}_{dip}$. In terms of a representation in which $\mathcal{H}_0$ is diagonal, the $g(\omega)$ embodies the effect of the diagonal and semidiagonal parts of $\mathcal{H}_{dip}$, which rearrange the entire dipole lattice in such a manner as to conserve over-all energy.

Two prescriptions have been used to calculate $g(\omega)$. The "exact" method evaluates moments of $g(\omega)$ by means of the usual trace formulas.[1,3] The "approximate" method considers $g(\omega)$ as the convolution of the resonance lines between which cross relaxation occurs. For instance, let $g_1(\omega)$ describe the line shape for absorption between states A and B, and let $g_2(\omega)$ describe the absorption line shape involving states C and D; then the $g(\omega)$, for a process in which one spin flips from A to B while another flips from D to C would be given by

$$g(\omega)=\int_{-\infty}^{\infty}g_1(\omega')g_2(\omega-\omega')d\omega', \tag{3}$$

where ω corresponds to the "leftover" energy, $\Delta E_{AB}-\Delta E_{CD}$. Since g_1 and g_2 are quite sharply peaked around $\omega=0$, g will also be peaked at $\omega=0$ in agreement with experiment.[4–8]

A number of difficulties have arisen, however, with the more detailed quantitative development of the theory. Among these we may mention the following:

(1) The "exact" moment calculations of Kopvillem,[9] Kiel,[10] and Hirono[11,12] indicate that the shape function which seems consistent with these calculations does not resemble the shape function which emerges from the "approximate" method. For 1- and 2-spin processes, Kopvillem predicts a square shape with a width of the same order as the resonant frequency, and essentially independent of concentration. Similarly, for a 3-spin process, Kiel and Hirono obtain leading terms for $g(\omega)$ which are in first-order concentration and frequency-independent. Equation (3), on the other hand, predicts a sharply peaked resonance function with a width and a concentration dependence similar to the absorption signal.

(2) The relationship of experimentally observed features to the predictions of either method seem somewhat unclear. For a concentrated system like LiF,[4] the convolution scheme provides a satisfactory interpretation for much, but not all, of the data. For dilute systems like ruby, the effects are also sharply resonant[6] and, at least for two-spin processes, appear to have a concentration-dependent width. On the other hand, this width seems to exceed considerably the absorption width expected from the convolution scheme. Furthermore, the intensity of the effect shows a concentration dependence unaccounted for by any calculation.

(3) There are also two conceptual difficulties. One

* This work was supported in part by the U. S. Army Signal Corps. under Contract DA-36-039-sc-89169.

[1] N. Bloembergen, S. Shapiro, P. S. Pershan, J. O. Artman, Phys. Rev. **114**, 445 (1958).
[2] J. A. Giordmaine, L. E. Alsop, F. R. Nash, and C. H. Townes, Phys. Rev. **109**, 302 (1958).
[3] J. H. Van Vleck, Phys. Rev. **74**, 1168 (1948).
[4] P. S. Pershan, Phys. Rev. **117**, 109 (1960).
[5] K. D. Bowers and W. B. Mims, Phys. Rev. **115**, 285 (1959).
[6] W. B. Mims and J. D. McGee, Phys. Rev. **119**, 1233 (1960).
[7] R. W. Roberts, J. H. Burgess, and H. D. Tenney, Phys. Rev. **121**, 997 (1961).
[8] A. A. Manenkov and A. M. Prokhorov, Zh. Experim. i Teor. Fiz. **42**, 75 (1962) [English transl.: Soviet Phys.—JETP **15**, 54 (1962)].
[9] U. K. Kopvillem, Fiz. Tverd. Tela. **2**, 1829 (1960) [English transl.: Soviet Phys.—Solid State **2**, 1653 (1960)].
[10] A. Kiel, Phys. Rev. **120**, 137 (1960).
[11] M. Hirono, J. Phys. Soc. Japan **16**, 766 (1961).
[12] M. Hirono, J. Radio Res. Labs. Japan **9**, 73 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

involves the sometimes ambiguous relationship of a function to its moments (to be more fully discussed in a future publication); the other concerns the precise origin and meaning of the shape function $g(\omega)$. Clarification of this latter question is one of the main objectives of the present work.

We shall re-examine the physical premises of the theory, construct a new mathematical formalism, and derive closed analytic expressions for the transition probability. While our theory is applicable to general cross-relaxation processes, we begin with the simplest possible process, in which a single spin flips.

II. ELEMENTARY PHYSICAL CONSIDERATIONS

We consider the Hamiltonian of Eq. (1).

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_{\rm dip}.$$

Without the interaction $\mathcal{H}_{\rm dip}$, each particle is independent, and its energy has eigenvalues $E_A{}^0$, $E_B{}^0$, etc. The eigenvalues of $\mathcal{H}$ are then simply sums of the single-particle eigenvalues. Because the particles are coupled by $\mathcal{H}_{\rm dip}$, their individual energy is no longer a constant of the motion. Nevertheless, as long as $\mathcal{H}_{\rm dip}$ is small, it is still convenient to choose the eigenstates of $\mathcal{H}_0$ as the zeroth-order eigenstates of $\mathcal{H}$. The diagonal and off-diagonal elements of $\mathcal{H}_{\rm dip}$ are then defined with respect to a representation in which $\mathcal{H}_0$ is diagonal. In first order, the diagonal elements of $\mathcal{H}_{\rm dip}$ can be thought of as shifting the energy of each particle, by an amount $E^{\rm dip}$, from its unperturbed value E^0. The energy shift will be different for each ion, because each ion sees a different local field. The energies of all the particles in a given Zeeman state will form a distribution centered on the Zeeman energy. In this sense, one speaks of the dipole interaction as "smearing" the Zeeman levels. The energy shift of an individual ion will depend not only on the specific environment in which the ion finds itself, but also on the Zeeman state in which it happens to be, since the diagonal elements of $\mathcal{H}_{\rm dip}$ are not all equal.

In higher order, transitions between states can be considered as induced by the "rotating" off-diagonal dipole matrix elements. Just as each ion sees a different diagonal perturbation, so it sees a different transition operator. Both are effects of the same local field.

We wish to calculate the probability that a particle will make a nonradiative transition from Zeeman state A, with unperturbed energy $E_A{}^0$, to Zeeman state B, with unperturbed energy $E_B{}^0$. Such a transition is possible only for a spin which happens to see a dipolar configuration such that, by flipping, it automatically changes the dipolar energy by the right amount to compensate for the change in Zeeman energy. Energy conserving transitions between separated Zeeman states A and B are possible because (1) the actual energy of a spin in state A or B will, in general, deviate from $E_A{}^0$ or $E_B{}^0$; (2) the magnitude of this perturbation will be different when the spin finds itself in state A and when in state B.

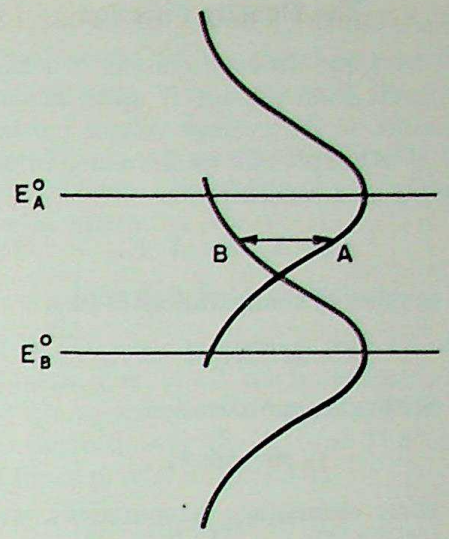


FIG. 1. Energy conserving transitions between different Zeeman states.

Figure 1 shows a distribution of states around Zeeman energies $E_A{}^0$ and $E_B{}^0$. We are interested in transitions between points like A and B. Such transitions are nonradiative, since they fulfill the requirement

$$E_A{}^{\rm total}=E_B{}^{\rm total}, \tag{4}$$

where

$$E^{\rm total}=E_0+E^{\rm dip}. \tag{5}$$

At the same time, such transitions change the relative populations of levels A and B, as well as the total magnetization of the system.

We do not require that each spin flip be accompanied by a dynamic rearrangement of the entire lattice to conserve energy, and in this respect our picture seems somewhat simpler than the one suggested by Bloembergen.

Stated in the present terms, the crux of the problem is to pick out those dipole configurations which meet the rather stringent requirement of Eq. (4). For each such configuration, we calculate the transition probability and then sum the result. The probability of finding a suitable dipole configuration can be considered as a density function. The total transition probability can then be interpreted as the average of the transition operator over this density function.

III. GENERAL EXPRESSION FOR CROSS-RELATION PROBABILITY

We now derive in integral form a general expression for the probability of a dipole-induced single-spin flip. For clarity's sake, we somewhat arbitrarily divide the discussion into three steps. First, we set up a formalism to project the suitable dipole configurations. Then we calculate the transition probability, summed over these configurations. Finally, we discuss the significance of the resulting formula from some alternate viewpoints.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The Density Function

We first calculate the total number of configurations that can result from putting N spins randomly into M lattice sites. If the average volume per lattice site is v, then M is simply the total volume in units of v. We define the spin concentration

$$n=N/M\equiv vN/V. \tag{6}$$

The total number of configurations is then

$$C=M!/N!(M-N)!. \tag{7}$$

By use of Stirling's approximation

$$x!=x^xe^{-x}, \tag{8}$$

and after some elementary computations, we obtain, for small n,

$$C=(1/N!)[M(1-n)]^N. \tag{9}$$

The factor $(1-n)$ accounts for the fact that we allow only one spin per lattice site rather than an arbitrary number. Stated in terms of the volume V

$$C=\frac{1}{N!}\left(\frac{1-n}{v}\right)^N V^N\equiv\frac{1}{N!}\left(\frac{1-n}{v}\right)^N\int d\mathbf{r}_1\cdots d\mathbf{r}_N. \tag{10}$$

Let us consider the simple question: What is the probability that a spin in state A will have some specified energy $E_A{}^{\text{total}}$? Now

$$E_A{}^{\text{total}}=E_A{}^0+E^{\text{dip}}. \tag{5}$$

$E_A{}^{\text{total}}$ and $E_A{}^0$ are numbers, but $E_A{}^{\text{dip}}$ is a function of the coordinates of every spin in the system. When we count the number of configurations, $C(E_A{}^{\text{total}})$, which fulfills the requirement (5), we proceed as before, except that we need to introduce the constraint [$\delta(x-x_0)$ is the Dirac delta function]

$$\delta[E_A{}^{\text{total}}-(E_A{}^0+E_A{}^{\text{dip}})]$$

into the volume integrals of Eq. (10):

$$C(E_A{}^{\text{total}})=\frac{1}{N!}\left(\frac{1-n}{v}\right)^N\int d\mathbf{r}_1\cdots d\mathbf{r}_N \times\delta[E_A{}^{\text{total}}-(E_A{}^0+E_A{}^{\text{dip}})]. \tag{11}$$

The fractional number of configurations that meet this constraint is evidently

$$f(E_A{}^{\text{total}})=C(E_A{}^{\text{total}})/C=\frac{1}{V^N}\int d\mathbf{r}_1\cdots d\mathbf{r}_N \times\delta[E_A{}^{\text{total}}-(E_A{}^0+E_A{}^{\text{dip}})]. \tag{12}$$

In similar fashion, the condition that a particle in state B has an energy $E_B{}^{\text{total}}$ is expressed by the constraint

$$\delta[E_B{}^{\text{total}}-(E_B{}^0+E_B{}^{\text{dip}})].$$

The condition, $E_A{}^{\text{total}}=E_B{}^{\text{total}}$, is expressed by the constraint

$$\delta[E_A{}^{\text{total}}-E_B{}^{\text{total}}].$$

Finally, to find the total number of "overlapping" states, one must integrate over-all values of $E_A{}^{\text{total}}$ and $E_B{}^{\text{total}}$. Thus, the fraction of configurations that satisfy our requirements is given by

$$f=\frac{1}{V^N}\int d\mathbf{r}_1\cdots d\mathbf{r}_N\delta[E_A{}^{\text{total}}-(E_A{}^0+E_A{}^{\text{dip}})] \times\delta[E_B{}^{\text{total}}-(E_B{}^0+E_B{}^{\text{dip}})] \times\delta[E_A{}^{\text{total}}-E_B{}^{\text{total}}]dE_A{}^{\text{total}}dE_B{}^{\text{total}}. \tag{13}$$

Using the elementary properties of δ functions,[13] and the notation

$$\hbar\omega=E_A{}^0-E_B{}^0, \tag{14}$$

$$\hbar\Delta\omega=E_B{}^{\text{dip}}-E_A{}^{\text{dip}}, \tag{15}$$

the expression (13) can be written more simply

$$f(\omega)=\frac{1}{V^N}\int d\mathbf{r}_1\cdots d\mathbf{r}_N\delta[\omega-\Delta\omega]. \tag{16}$$

Replacing the δ function by its Fourier transform, we have

$$f(\omega)=\frac{1}{V^N}\int d\mathbf{r}_1\cdots d\mathbf{r}_N\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\rho(\omega-\Delta\omega)}d\rho. \tag{17}$$

The $\Delta\omega$ in Eq. (17) is a function of all the $\mathbf{r}$'s. We assume that the contributions of all the particles to the energy shift $\Delta\omega$ are additive:

$$\Delta\omega=\sum_i\Delta\omega(\mathbf{r}_i) \tag{18}$$

$$\equiv\frac{1}{\hbar}\sum_i\mathfrak{IC}_{BB}{}^{\text{dip}}(\mathbf{r}_i)-\mathfrak{IC}_{AA}{}^{\text{dip}}(\mathbf{r}_i). \tag{19}$$

The approximation of additivity Eq. (18) involves an error of the order of the ratio of the dipolar energy to the Zeeman energy, which is small except for very low magnetic fields. With Eq. (18), the $\mathbf{r}$ integrals implied in Eq. (17) are now separable, and we may write

$$f(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(\rho)e^{-i\rho\omega}d\rho \tag{20}$$

with

$$\Phi(\rho)=\lim_{N\to\infty}\left[\frac{1}{V}\int d\mathbf{r}e^{i\rho\Delta\omega(\mathbf{r})}\right]^N. \tag{21}$$

The function $\Phi(\rho)$ and its Fourier transform $f(\omega)$ have been discussed by many authors.[14–17]

[13] H. Heitler, *Quantum Theory of Radiation* (Oxford University Press, New York, 1957), 3rd ed., p. 66 ff.
[14] H. Margenau and J. Watson, Rev. Mod. Phys. 8, 22 (1936).
[15] H. Margenau, Phys. Rev. 48, 755 (1935).
[16] P. W. Anderson, Phys. Rev. 76, 647 (1949).
[17] P. W. Anderson, Phys. Rev. 82, 342 (1951).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We are thus able to calculate, in terms of known functions, the fraction of configurations for which an *energy conserving* transition between two Zeeman states is possible. This fraction is a function of the energy separation $\hbar\omega$ of the Zeeman states. We have referred to it as a density function because it gives the density of available configurations as a function of ω.

The Transition Operator

The introduction of the transition operator is straightforward. This operator is $(\sum M_i)^2$, where M_i is the off-diagonal dipole matrix element connecting Zeeman states A and B, and involving the ith spin and the reference spin we are considering. The M_i are functions of the coordinate variables. Consequently, the transition matrix element is different for each allowed configuration of spins. This means that in the process of picking out suitable configurations from the ensemble, we weight each configuration separately with its own transition operator. In other words, the transition matrix element goes inside the coordinate integrals of Eq. (13). Analogous to Eq. (16), we now obtain for the actual transition probability

$$W(\omega)=\frac{2\pi}{\hbar}\frac{1}{V^N}\int d\mathbf{r}_1\cdots d\mathbf{r}_N[\sum_i M^2(\mathbf{r}_i)]\delta(\omega-\Delta\omega). \quad (22)$$

We note that

$$(\sum M_i)^2=\sum M_i^2. \quad (23)$$

The cross terms vanish in the understood integration over time because of the different time dependence of each M_i.

Each term in the sum over i gives an identical integral, so that

$$W(\omega)=\frac{2\pi}{\hbar}\frac{n}{v}\int_{-\infty}^{\infty}\chi(\rho)\Phi(\rho)e^{-i\rho\omega}d\rho, \quad (24)$$

where

$$\chi(\rho)=\frac{1}{2\pi}\int d\mathbf{r}M^2(\mathbf{r})e^{i\rho\Delta\omega(\mathbf{r})}, \quad (25)$$

and $\Phi(\rho)$ is defined in Eq. (21). Actually, the N in Eq. (21) is here replaced by $N-1$, but this makes no difference whatsoever since

$$\lim_{V\to\infty}\left[\frac{1}{V}\int d\mathbf{r}e^{i\rho\Delta\omega(\mathbf{r})}\right]=1. \quad (26)$$

We remark parenthetically that our discussion is applicable rather more simply to resonance lines. The differences are the following: Our constraint $\Delta E^{\text{total}}=0$ is replaced by the constraint $\Delta E^{\text{total}}=\Delta E^{\text{Zeeman}}$, which merely shifts the center frequency. The transition operator is changed to S_x^2, without any coordinate dependence, which eliminates $\chi(\rho)$ in Eq. (24). In particular, $\Phi(\rho)$ remains the same. Thus, $\Phi(\rho)$ is essentially the Fourier transform of the resonance line shape. The function $\Phi(\rho)$ is therefore accessible to us through experiment, even if we were not able to calculate it.

Equation (24) is our general formula for the single-spin cross-relaxation probability, as a function of the Zeeman level separation, $\hbar\omega$.

Alternate Interpretations

Equation (24) can be given a variety of physical interpretations. First of all, the function $\chi(\rho)$ is closely related to the autocorrelation function of the transition operator, summed over the lattice. The autocorrelation of a function $F(t)$ is defined as

$$C(\rho)=\langle F(t+\rho)F^*(t)\rangle_{\text{time average}}. \quad (27)$$

If F is an operator, then in the Heisenberg representation

$$F(t+\rho)=\exp(i\mathcal{H}\rho/\hbar)F(t)\exp(-i\mathcal{H}\rho/\hbar). \quad (28)$$

When F and $\mathcal{H}$ refer to the off-diagonal and diagonal parts of $\mathcal{H}^{\text{dip}}$, we obtain

$$M(t+\rho)=e^{i\rho\Delta\omega}M(t) \quad (29)$$

and

$$C(\rho)=\langle|M(t)|^2\rangle_{\text{time average}}e^{i\rho\Delta\omega}. \quad (30)$$

Summing $C(\rho)$ over the lattice yields our $\chi(\rho)$. We have defined, in Eq. (27), an unnormalized autocorrelation function, which is not usual. In point of fact, our $\chi(\rho)$ is unnormalized. It accounts for the magnitude of the transition matrix element as well. Indeed

$$\int_{-\infty}^{\infty}W(\omega)d\omega=\frac{2\pi}{\hbar}\frac{n}{v}\int d\mathbf{r}M^2. \quad (31)$$

This can be seen immediately by evaluating the Fourier transform of $W(\omega)$, that is, $\chi(\rho)\Phi(\rho)$, at $\rho=0$.

If $\chi(\rho)$ is the autocorrelation function of the transition operator, then its Fourier transform is the power spectrum of this operator. The Fourier transform of the function Φ, as we have seen, is essentially the resonance line shape. The transition probability is thus given by the scaled power spectrum of the transition operator convoluted with the spectrum of the broadened levels. If the transition operator were a δ function in ω space, then we would surely expect that energy could not be conserved in a single-spin flip unless the Zeeman level separation were of the order of the dipolar broadening. This assumption is implicit in the calculation of the cross-relaxation probability according to the scheme represented in Eq. (3). Conversely, if the levels were sharp, they would see a resonant component of the interaction mechanism only when the level separation fell within the Fourier spectrum of this mechanism. This latter consideration, in fact, provided the basis for the calculation of spin-lattice relaxation in Bloembergen, Purcell, and Pound's classic

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

nuclear relaxation paper.[18] In our case, the levels as well as the transition operator have a spectrum, and the resulting width is obtained by convolution.

Alternatively, one can interpret Eq. (24) simply as an application of the "Golden Rule":

$$W=(2\pi/\hbar)|\mathfrak{IC}_{\text{transition}}|^2\rho(E). \tag{32}$$

We must be somewhat careful in our interpretation of this formula. Any particular single spin, in flipping from Zeeman level A to level B, does not have available a continuum of states, but exactly one well-defined, albeit perturbed state. This state is either degenerate with the initial state or it is not; accordingly, $\rho(E)$ is either 1 or 0. We must then sum these Kronecker-δ probabilities over the discrete manifold of N spin states. If we replace the discrete manifold by a continuum of levels, and the Kronecker δ by a Dirac δ, we are immediately led to the equations that we have already derived.

In a third interpretation, we could regard $\rho(E)$ in Eq. (32) as referring to the lattice as a whole. It would then be a true density function. Given two Zeeman states separated in unperturbed energy by ΔE^0, and broadened into a continuous set of states by the dipole interaction, then ρ would represent the average occupancy of that subset of states for which $\Delta E^{\text{total}}=0$. Thus, $\rho(E)$ would correspond to our $f(\omega)$ in Eq. (21). When averaging the transition matrix element over the distribution ρ, one must evidently integrate over all the random variables involved. In the present case these include the space coordinates. This is precisely what we have done. This interpretation is implicit in the spin-spin line shape formula first given by Waller,[19] and quoted in Ref. 18. We note that in Ref. 1, however, the spatial average of the matrix element and the level density function are calculated separately.

We could indeed have written down Eqs. (24) and (25) almost immediately by invoking any of these three considerations. Our more pedestrian presentation, while perhaps less elegant, has been aimed at giving as clear a justification as possible for the mathematical expression of the theory.

Finally, we remark that the concentration dependence of the shape of $W(\omega)$ resides entirely in Φ, not in χ. The level broadening described by Φ depends linearly on the concentration for low concentrations and is proportional to the square root of the concentration for high concentrations, if the broadening is due to the dipole field. This, of course, is well known. The power spectrum of the transition operator, described by χ depends only upon the spatial integrals of the relevant coordinate functions. The *relative* occupation probabilities of the lattice sites are clearly not affected by the concentration; hence, the functional form of these integrals is not affected. Of course, the concentration enters also as a scale factor in Eq. (24) governing the absolute magnitude of $W(\omega)$.

[18] N. Bloembergen, E. M. Purcell, and R. V. Pound, Phys. Rev. 73, 679 (1947).
[19] I. Waller, Z. Physik 79, 370 (1932).

IV. EVALUATION OF THE INTEGRALS

Our next task is to evaluate the function χ and its Fourier transform. The transition probability itself will be proportional to the convolution of this transform with the transform of Φ.

Definition of $\Delta\omega$ and M

As a preliminary, we must define the functions represented by $\Delta\omega(\mathbf{r})$ and $M(\mathbf{r})$ in Eqs. (24) and (25). In Bloembergen's tabulation of dipole matrix elements,[1] his A and B functions correspond to our $\Delta\omega(\mathbf{r})$ and his C, D, E, and F functions correspond to our $M(\mathbf{r})$. Abbreviating his notation, we write

$$\Delta\omega(\mathbf{r})=q(3\cos^2\theta-1)/r^3, \tag{33}$$

$$M(\mathbf{r})=g(\sin\theta\cos\theta)/r^3, \quad \text{if} \quad \Delta m=\pm1, \tag{34}$$

$$M(\mathbf{r})=g(\sin^2\theta)/r^3, \quad \text{if} \quad \Delta m=\pm2. \tag{35}$$

Here q and g contain appropriate diagonal and off-diagonal spin-matrix elements. These expressions are valid only for two special cases: (1) There is no crystal field, or, (2) if there is a crystal field, the magnetic axis coincides with the crystal axis. In the more general case, when the magnetic field makes an arbitrary angle with the crystal axis, the expressions are generalized as follows:

$$\Delta\omega(\mathbf{r})=\sum_{k=0}^{2} q_k P_{2k}\equiv\frac{1}{r^3}f(\theta), \tag{36}$$

$$M(\mathbf{r})=\frac{1}{r^3}\sum_{k=0}^{2} g_k P_{2k}\equiv\frac{1}{r^3}m(\theta). \tag{37}$$

Here P_{2k} are the Legendre polynomials P_{20}, P_{21}, P_{22}, and the q_k and g_k now represent exceedingly complicated functions of the spin-matrix elements, the magnetic field, and the crystal field parameters. But even in the most general case, $\Delta\omega$ and M resolve into a product of an angular function and r^{-3}.

Evaluation of χ

Instead of evaluating $\chi(\rho)$ directly, we consider first a slightly more general integral than the one in Eq. (25):

$$I(\rho,n)=\int_{-1}^{1} d\cos\theta\int_{r_0}^{\infty} r^2dr\frac{m^2(\theta)}{r^{3n}}e^{[i\rho f(\theta)/r^3]}, \tag{38}$$

so that

$$\chi(\rho)=I(\rho,2). \tag{39}$$

The lower limit of the r integral is the nearest-neighbor distance r_0. This distance can never be zero since two particles cannot occupy the same spot.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

To perform the integration over r, we put

$$a=\rho f(\theta), \tag{40}$$

$$t=-ia/r^3. \tag{41}$$

We obtain

$$I(\rho,n)=\int_{-1}^{1} d\cos\theta \frac{m^2(\theta)}{3(-ia)^{n-1}}\int_0^{-ia/r_0^3} e^{-t}t^{n-2}dt. \tag{42}$$

The t integral will be recognized as defining the incomplete γ function[20] so that

$$I(\rho,n)=\int_{-1}^{1} d\cos\theta\left(\frac{m^2(\theta)}{3r_0^{3(n-1)}}\right)\frac{\gamma(n-1,\ -ia/r_0^3)}{(-ia/r_0^3)^{n-1}}, \tag{43}$$

$$\chi(\rho)=\int_{-1}^{1} d\cos\theta\left(\frac{m^2(\theta)}{3r_0^3}\right)\frac{\gamma(1,\ -ia/r_0^3)}{-ia/r_0^3}. \tag{44}$$

After a slight manipulation of the expansion[20]

$$\gamma(a,x)=\sum_{n=2}^{\infty}\frac{(-)^n x^{a+n}}{n!(a+n)}, \tag{45}$$

the result of Eq. (44) may be written

$$\chi(\rho)=\int_{-1}^{1} d\cos\theta\left(\frac{m^2(\theta)}{3r_0^3}\right)\left(\frac{1-e^{i\rho f(\theta)/r_0^3}}{-i\rho f(\theta)/r_0^3}\right). \tag{46}$$

The result (46) could, of course, be obtained directly from (38), but we shall need the general expression (43) later on.

We must still do the angular integral and then take the Fourier transform. These operations cannot be performed analytically for the most general angular functions $m(\theta)$ and $f(\theta)$ defined in Eqs. (36) and (37). We shall, therefore, consider a number of special cases, amenable to calculation, from which the significant features of the result will become apparent.

Let us assume, as is invariably the case for a transition probability $W(\omega)$ that is symmetric in ω, that the diagonal perturbations $\Delta\omega$ and $-\Delta\omega$ occur in pairs, the perturbation of either sign being associated with the same off-diagonal M. Under this assumption, we shall make, in turn, a sequence of approximations:

Case a. We replace both angular functions in Eq. (46) by some constant average value, $\langle m^2\rangle$ and $\langle f\rangle$.

Case b. We replace $m^2(\theta)$ by some average value; replace $f(\theta)$ by a function defined only by the property that, as $\cos\theta$ goes from -1 to $+1$, the function assumes with equal likelihood all values from $-L$ to L.

Case c. We replace $m(\theta)$ by $g(\sin\theta\cos\theta)$ and $f(\theta)$ by $q(3\cos^2\theta-1)$. These functions can be considered as specializations of Eqs. (36) and (37) by setting $q_1=q_2=g_0=g_2=0$. They are the physically correct functions when the magnetic field is parallel to the c axis.

Case d. Finally, we consider the possibility that $\Delta\omega$ does not occur in pairs, so that $W(\omega)$ will not be symmetric. We shall evaluate the integrals with the physically exact functions $m(\theta)=g(\sin\theta\cos\theta)$ and $f(\theta)=q(3\cos^2\theta-1)$.

The actual computations are quite involved. They are outlined in the Appendix. Here we only present and discuss the results.

It is convenient to define the quantity

$$\epsilon=1/r_0^3. \tag{47}$$

Then expressions like $q\epsilon$, $L\epsilon$, $\langle f\rangle\epsilon$ all represent, at least to order of magnitude, the effective nearest-neighbor dipole energy.

Case a

$$\chi(\omega)=\frac{4\pi\langle m^2\rangle}{3r_0^3\epsilon\langle f\rangle}\left[u_1\left(\frac{\omega}{\epsilon\langle f\rangle}+1\right)-u_1\left(\frac{\omega}{\epsilon\langle f\rangle}-1\right)\right], \tag{48}$$

$$\begin{aligned} u_1(x)&=0, \quad x<0\\ &=1, \quad x>0. \end{aligned} \tag{49}$$

Equation (48) represents a square shape of range $-\epsilon\langle f\rangle<\omega<\langle f\rangle\epsilon$, and magnitude $4\pi\langle m^2\rangle/3r_0^3\epsilon\langle f\rangle$. The area is $8\pi\langle m^2\rangle/3r_0^3$. In ruby, the nearest-neighbor dipole energy in the absence of exchange would be about 2.3 kMc/sec, so that this result is rather reminiscent of Kopvillem's.[8]

What happens to this disconcerting square shape, once we assume almost any kind of nonconstant angular function $f(\theta)$, is intuitively apparent. As we integrate over θ we shall be summing different square waves corresponding to different values of f. These rectangles are all centered at the origin and their height is inversely proportional to their extent along the ω axis. The result must inevitably be a function with a maximum at the origin and a falling monotonically to zero in the wings. If the range of $f(\theta)$ includes the value zero, which it does in all practical cases, the resulting function must approach infinity at the origin. The area of $\chi(\omega)$ is always equal to $\chi(\rho)_{\rho=0}$, and inspection of Eq. (46) shows that this area still remains $8\pi m^2/3r_0^3$, regardless of any assumption about m and f. Consequently, the convolution of this singular function with the bell-shaped transform of $\Phi(\rho)$ will always give a finite, more or less bell-shaped result. We already know, then, from these very general considerations, that the cross-relaxation probability $W(\omega)$ must have a resonant shape.

Case b

$$\begin{aligned} \chi(\omega)&=\frac{8\pi\langle m^2\rangle}{3r_0^3}\frac{1}{2L\epsilon}\ln\frac{L\epsilon}{\omega}, \quad |\omega|<L\epsilon\\ &=0, \quad |\omega|>L\epsilon. \end{aligned} \tag{50}$$

[20] A. Erdely, F. Oberhettinger, W. Magnus, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1954), Chap. 6.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

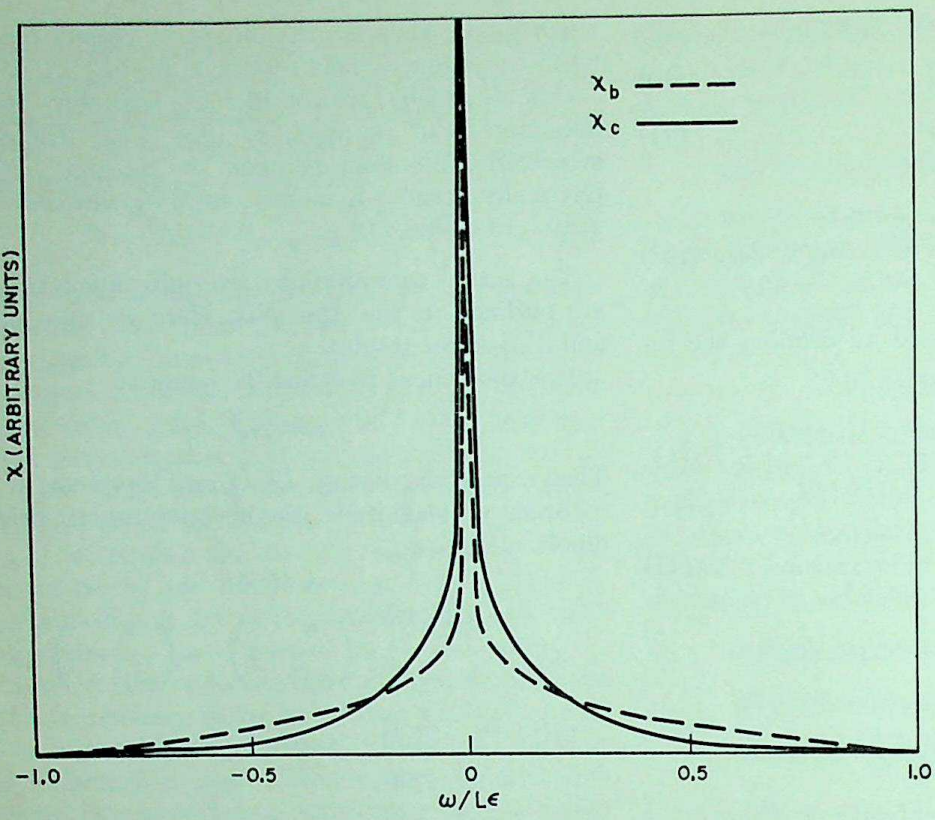


FIG. 2. χ_b [Eq. (50)] and χ_c [Eq. (53)].

This function has a logarithmic singularity at the origin and an area of $8\pi\langle m^2\rangle/3r_0^3$, as expected. The "half-power" width is, of course, undefined. If we define a half-width $\omega_{1/2}$ by the relation

$$\int_0^{\omega_{1/2}} \chi(\omega)d\omega = \tfrac{1}{2}\int_0^{\infty} \chi(\omega)d\omega, \tag{51}$$

then $\omega_{1/2}$ turns out to be

$$\omega_{1/2} = L\epsilon/5.4 \cong L\epsilon/2e. \tag{52}$$

Case c

If $|\omega| < q\epsilon$,

$$\chi(\omega) = \frac{4\pi g^2}{27\sqrt{3}r_0^3 q\epsilon}\left\{\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1} -\tfrac{1}{3}\left[\left(1-\frac{\omega}{q\epsilon}\right)^{1/2}\left(5+\frac{\omega}{q\epsilon}\right)+\left(1+\frac{\omega}{q\epsilon}\right)^{1/2}\left(5-\frac{\omega}{q\epsilon}\right)\right] +\ln\left|\frac{[1-(\omega/q\epsilon)]^{1/2}+1}{[1-(\omega/q\epsilon)]^{1/2}-1}\right|+\ln\left|\frac{[1+(\omega/q\epsilon)]^{1/2}+1}{[1+(\omega/q\epsilon)]^{1/2}-1}\right|\right\}. \tag{53a}$$

If $q\epsilon < |\omega| < 2q\epsilon$,

$$\chi(\omega) = \frac{4\pi g^2}{27\sqrt{3}r_0^3 q\epsilon}\left\{\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1} -\frac{1}{3}\left(1+\frac{|\omega|}{q\epsilon}\right)^{1/2} \times\left(5-\frac{|\omega|}{q\epsilon}\right)+\ln\left|\frac{[1+(|\omega|/q\epsilon)]^{1/2}+1}{[1-(|\omega|/q\epsilon)]^{1/2}-1}\right|\right\}. \tag{53b}$$

If $|\omega| > 2q\epsilon$,

$$\chi(\omega) = 0. \tag{53c}$$

In Fig. 2 we display this function, together with the simple logarithmic function of Eq. (50). In Eq. (50) we have chosen $L=2$ and $\langle m^2\rangle = \langle \sin^2\theta \cos^2\theta\rangle = 2/15$, and in Eq. (53) we have put $g=q=1$, so that the parameters match in both expressions. With this choice of numbers, the logarithmic function behaves like $\ln|\omega|^{-0.0444/\epsilon}$ near the origin, whereas the more complicated function behaves like $\ln|\omega|^{-0.0856/\epsilon}$. The exact function has the sharper singularity. The areas, as well as all qualitative features are the same.

Case d

If $\omega < -q\epsilon$,

$$\chi(\omega) = 0. \tag{54a}$$

If $-q\epsilon < \omega < 0$,

$$\chi(\omega) = \frac{4\pi g^2}{27\sqrt{3}r_0^3 q\epsilon}\left\{-\frac{1}{3}\left(\frac{\omega}{q\epsilon}+1\right)^{1/2}\left(\frac{\omega}{q\epsilon}-5\right) +\ln\left|\frac{[1+(\omega/q\epsilon)]^{1/2}+1}{[1+(\omega/q\epsilon)]^{1/2}-1}\right|\right\}. \tag{54b}$$

If $0 < \omega < 2q\epsilon$,

$$\chi(\omega) = \frac{4\pi g^2}{27\sqrt{3}r_0^3 q\epsilon}\left\{\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1} -\frac{1}{3}\left(\frac{\omega}{q\epsilon}+1\right)^{1/2} \times\left(\frac{\omega}{q\epsilon}-5\right)+\ln\left|\frac{[1+(\omega/q\epsilon)]^{1/2}+1}{[1+(\omega/q\epsilon)]^{1/2}-1}\right|\right\}. \tag{54c}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

If $\omega > 2q\epsilon$,

$$\chi(\omega) = 0. \tag{54d}$$

We remark that this function is not symmetric in ω. The asymmetry is important primarily in the wings. $\chi(\omega)$ and $\chi(-\omega)$ assume identical form as ω approaches zero. The associated function $\Phi(\omega)$ behaves similarly. It seems likely, therefore, that any asymmetry arising from the present considerations would be accessible only to measurements made under conditions where large amounts of energy are transferred to or from the dipole system.

Qualitatively there is little difference between the symmetric case c and the nonsymmetric case d, just as little difference results from using the approximate f and m of case b instead of the f and m of case c, which represent the dipole interaction exactly provided the Zeeman and crystal field operators commute. In all cases there is a logarithmic singularity at the origin, and an over-all width determined by the effective nearest neighbor dipole energy. It is clear that these characteristics of χ can be expected to hold even when the most general form of f and m allowed by Eqs. (36) and (37) is physically appropriate.

Convolution

As indicated by Eq. (24), the function we seek, $W(\omega)$, is proportional to the convolution of $\Phi(\omega)$ and $\chi(\omega)$.

The widths of two convoluted functions, generally speaking, add in some fashion. Two Lorentz curves add width algebraically, two Gaussians add width rms wise. In particular, in the limit that the width of one function goes to zero, the resultant width approaches the width of the other function. A δ function simply leaves its partner intact. Our function $\chi(\omega)$ is an exception to this rule. Its peculiar properties under convolution are of considerable importance in determining the shape of $W(\omega)$.

As far as the effect on the resultant width is concerned, the exact shape of Φ does not make much difference, particularly when Φ is either very much broader or very much narrower than χ. If we use the simplified form of $\chi(\omega)$ given in Eq. (50), and simplify $\Phi(\omega)$ by a square shape of unit area and half-length a, we can do the convolution analytically.

If $a > L\epsilon$, it turns out that the resulting half-width is always identically a. Thus, when Φ is broader than χ, χ adds no width at all, but merely smoothes out the discontinuity of the square shape. In this case χ is practically equivalent to a δ function. The more interesting case is when $a < L\epsilon$. Then the convolution gives:

If $0 < |\omega| < (L\epsilon - a)$,

$$W(\omega) = \frac{C}{4}\left[\frac{2}{L\epsilon} + \frac{1}{L\epsilon}\left(\frac{\omega}{a}+1\right)\ln\left|\frac{L\epsilon}{\omega+a}\right| + \frac{1}{L\epsilon}\left(1-\frac{\omega}{a}\right)\ln\left|\frac{L\epsilon}{a-\omega}\right|\right]. \tag{55a}$$

If $(L\epsilon - a) < |\omega| < (L\epsilon + a)$,

$$W(\omega) = \frac{C}{4}\left[\frac{1}{a} + \frac{1}{L\epsilon}\left(1-\frac{|\omega|}{a}\right)\left(\ln\left|\frac{L\epsilon}{a-|\omega|}\right|+1\right)\right]. \tag{55b}$$

If $|\omega| > (L\epsilon + a)$,

$$W(\omega) = 0. \tag{55c}$$

$$C = \frac{2\pi}{\hbar}\frac{n}{v}\frac{8\pi\langle m^2\rangle}{3r_0^3}. \tag{55d}$$

From Eq. (55a), the quantity $\omega_{1/2}$, the half-width of the resultant $W(\omega)$, can easily be calculated. We define

$$N = \omega_{1/2}/a. \tag{56}$$

Then Eq. (55a) can be written

$$W(0) = \frac{C}{2L\epsilon}\left[1+\ln\frac{L\epsilon}{a}\right], \tag{57}$$

$$W(Na) = \frac{C}{2L\epsilon}\left[1+\frac{(N+1)}{2}\ln\frac{L\epsilon}{(N+1)a} - \frac{(N-1)}{2}\ln\frac{L\epsilon}{(N-1)a}\right]. \tag{58}$$

If the χ width $L\epsilon/2e$ is much larger than a, $\omega_{1/2}$ will be likewise much larger than a, and N will be large. If we expand the logarithms to first power in $1/N$, Eq. (58) can be reduced to

$$W(Na) = \ln(L\epsilon/Na). \tag{59}$$

If we now solve $W(Na) = \frac{1}{2}W(0)$, we obtain

$$Na = \omega_{1/2} = [2a(L\epsilon/2e)]^{1/2}. \tag{60}$$

This is a remarkable result. As the Φ width a goes to zero, $\omega_{1/2}$ of W does not approach $\omega_{1/2}$ of χ, $L\epsilon/2e$, as one would intuitively expect, but rather it approaches zero. This strange property of multiplying widths rather than adding them turns out to be very nearly true even when Φ and χ have comparable widths, as can be seen from Fig. 3. We have done the convolution of the physically real χ of *Case c*, using square waves, Gaussians, and Lorentzians for Φ. The convolutions were done numerically on the IBM 7090, and the behavior of the half-widths is illustrated in Fig. 3.

The shape of the resulting function is also noteworthy. When Φ and χ have comparable widths, the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

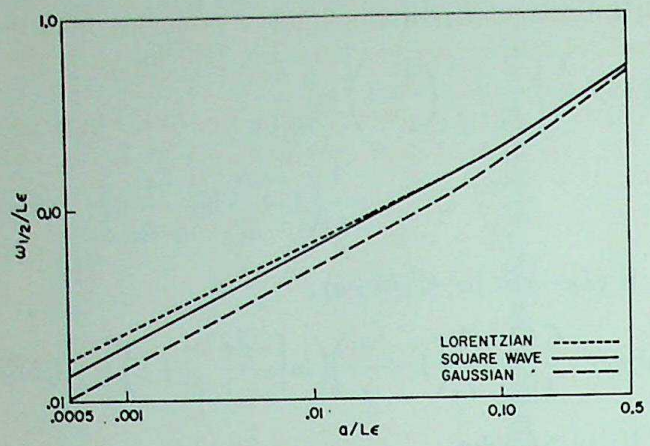


FIG. 3. Half-width ($\omega_{1/2}$) of convolution of Φ with χ_c, versus half-width (a) of Φ. For Φ = square wave, $\omega_{1/2} \sim a^{0.50}$; Φ = Gaussian, $\omega_{1/2} \sim a^{0.50}$, Φ = Lorentzian, $\omega_{1/2} \sim a^{0.48}$.

shape is nearly that of Φ. As Φ becomes narrower than χ, the shape of their convolution somewhat resembles an upside-down funnel. These changes in the shape of $W(\omega)$ are illustrated in Fig. 4. [The extreme changes in shape have necessitated drawing Fig. 4(a) on a different scale than Figs. 4(b) and 4(c), as indicated.]

The significance of the present result is that it is the key to the concentration dependence of the width of $W(\omega)$. As we have already remarked, this dependence is contained entirely in Φ. Our χ has a width of the order of the effective nearest-neighbor interaction. In general, one would expect χ to be considerably broader than Φ. If χ added width, in convolution, like normal functions, one would predict a concentration-dependent width only for very high concentrations. The unusual behavior of χ in this regard predicts a dependence proportional to the square root of the concentration, even at vanishing concentrations. Furthermore, χ introduces the multiplicative factor $(L\epsilon/e)^{1/2}$, which makes the widths substantially larger than the resonance width a. This convolution property of χ thus enables us to meet all the experimentally given requirements: to have $W(\omega)$ much narrower than either the resonant frequency or the maximum dipole energy, but much broader than the absorption width, *and* at the same time concentration-dependent.

Finally, it is clear from Figs. 2 and 3 that all the essential features of our results hold whether we use the χ of Eq. (50) or the more complicated one of Eq. (53). Thus, it is unnecessary to evaluate χ explicitly for general $\Delta\omega$ and M [Eqs. (36) and (37)]. The relevant parameters are the average of M^2, the range of f, which we have called L, and the value of r_0. These parameters can, in principle, be readily calculated for any physical situation, and they will define χ in terms of Eq. (50).

V. SUMMARY

We have laid a physical foundation for a general theory of cross relaxation, and we have applied our considerations particularly to single-spin transitions. Our formulation makes it unnecessary to appeal to moments or to plausibility arguments; rather, it yields the relevant functions exactly, within the limits of the physical assumptions. We review the principal steps of the argument.

(1) For each spin, the dipole interaction causes both diagonal, "static," energy shifts, and transitions between Zeeman levels.

(2) Due to the energy shifts, transitions between Zeeman states are possible which strictly conserve energy.

(3) The transition probability depends on the density of degenerate states. This density is obtained by counting the number of dipole configurations which meet the relevant energy constraint. The transition probability is given by integrating over this distribution the square of the off-diagonal dipole elements. We note a certain analogy to the method of the microcanonical ensemble, where the trick also is to calculate

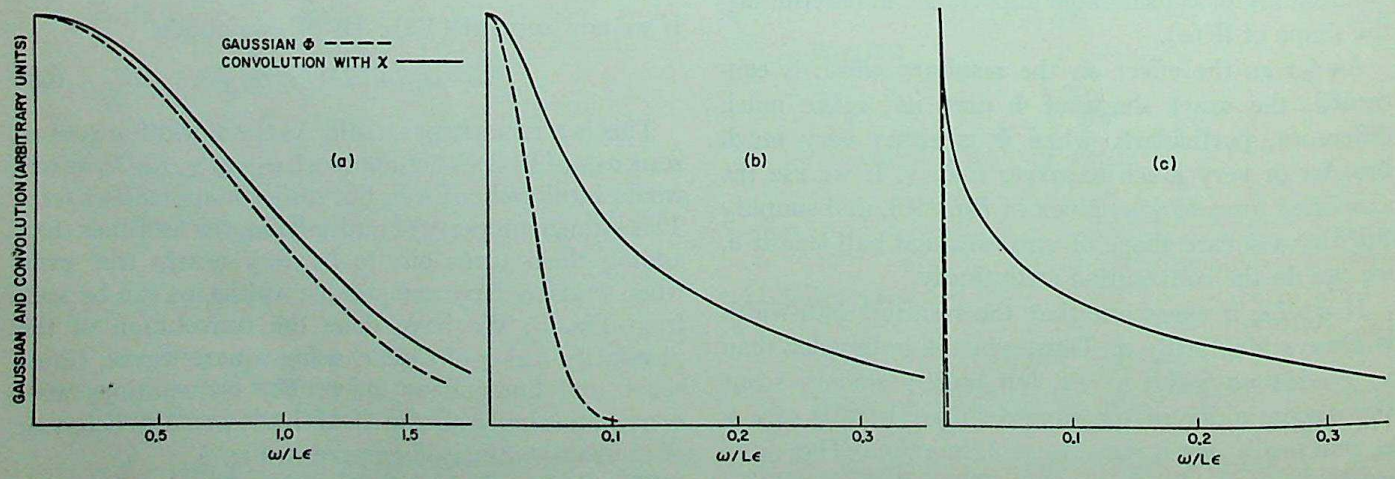


FIG. 4. Convolution of χ_c and Gaussian Φ, for various values of $\omega_{1/2}$ (Gaussian)/$\omega_{1/2}(\chi_c)$. (a) $\omega_{1/2}(G)/\omega_{1/2}(\chi_c) = 5$. (b) $\omega_{1/2}(G)/\omega_{1/2}(\chi_c) = 0.2$. (c) $\omega_{1/2}(G)/\omega_{1/2}(\chi_c) = 0.008$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the phase volume (number of configurations) corresponding to a given energy.

(4) We showed that the equations derived in this manner can also be derived from the Fourier analysis of the dipole operator and from any of several interpretations of first-order time-dependent perturbation theory.

(5) For single flips, the transition probability $W(\omega)$ was given by the convolution of two functions, $\chi(\omega)$ and $\Phi(\omega)$. χ is essentially the power spectrum of the off-diagonal elements of $\mathcal{H}^{\rm dip}$. To an excellent approximation, it depends on ω as $\ln|1/\omega|$. Φ describes essentially the level broadening due to the diagonal elements of $\mathcal{H}^{\rm dip}$. It is equivalent to the absorption line shape. The convolution of Φ and χ, which yields $W(\omega)$, has a sharp peak whose width is proportional to the geometric mean of the χ width and the Φ width. This sharp peak is superposed on a broad spectrum whose width is of the same order as the nearest-neighbor dipole energy. The transition probability per spin is linear in the concentration n, and for low concentrations its width in ω space is proportional to the square root of the concentration.

We have considered only single-spin flips explicitly, while experimental measurements are made on double, triple, or higher multiple processes. Nevertheless, the frequency and concentration dependence of our cross-relaxation probability satisfies a number of experimental features which have not been accounted for previously in terms of a single unified calculation.

The explicit extension of the theory to multiple-spin processes is the principal subject of the succeeding paper (part II); detailed quantitative application of the entire theory to typical experiments is made in part III.

APPENDIX: EVALUATION OF χ UNDER VARIOUS APPROXIMATIONS

We refer to Eq. (46). If we assume that the diagonal perturbations $\Delta\omega$ and $-\Delta\omega$ come in pairs, then Eq. (46) becomes

$$\chi(\rho)=\int_{-1}^{1} d\cos\theta\left(\frac{m^2(\theta)}{3r_0^3}\right)\left(\frac{e^{-i\rho f(\theta)\epsilon}-e^{i\rho f(\theta)\epsilon}}{-i\rho f(\theta)\epsilon}\right) \quad \text{(A1a)}$$

$$\equiv\int_{-1}^{1} d\cos\theta\left(\frac{2m^2(\theta)}{3r_0^3}\right)\frac{\sin[\rho f(\theta)\epsilon]}{\rho f(\theta)\epsilon}, \quad \text{(A1b)}$$

where ϵ is defined in Eq. (47).

Case a. If we set $m^2(\theta)=\langle m^2\rangle$ and $f(\theta)=\langle f\rangle$, then Eq. (A1b) reduces to

$$\chi(\rho)=(4m^2/3r_0^3)\sin\epsilon\langle f\rangle\rho/\epsilon\langle f\rangle\rho). \quad \text{(A2)}$$

The Fourier transform of $(\sin x)/x$ is well known.[21] It is a square wave of amplitude π extending from -1 to 1. Hence, we arrive at the Eq. (48) for $\chi(\omega)$, given in the text.

Case b. If we replace f by a function which goes from $-L$ to L as $\cos\theta$ goes from -1 to 1, Eq. (A1) becomes

$$\chi(\rho)=\int_{-L}^{L} df\left(\frac{2\langle m^2\rangle}{3Lr_0^3}\right)\frac{\sin\epsilon f\rho}{\epsilon f\rho}. \quad \text{(A3)}$$

The integral will be recognized as essentially the sine-integral function (Ref. 20, p. 145), and so

$$\chi(\rho)=\frac{4\langle m^2\rangle}{3r_0^3}\frac{\mathrm{Si}(L\epsilon\rho)}{L\epsilon\rho}. \quad \text{(A4)}$$

The Fourier transform of $\mathrm{Si}(x)/x$ is also well known,[21] and so we arrive at Eq. (50) in the text.

Case c. If we set $f=q(3\cos^2\theta-1)$ and $m^2=g^2\sin^2\theta\times\cos^2\theta$, Eq. (A1) becomes

$$\chi(\rho)=\frac{g^2}{3r_0^3\epsilon q}\int_{-1}^{1} d\cos\theta\left(\frac{\sin^2\theta\cos^2\theta}{3\cos^2\theta-1}\right)\left(\frac{1}{-i\rho}\right)\times\{\exp[-i\rho\epsilon q(3\cos^2\theta-1)]-\exp[i\rho\epsilon q(3\cos^2\theta-1)]\}. \quad \text{(A5)}$$

The angular integrals in Eq. (A5) cannot be done analytically as they stand. Instead, we shall avail ourselves of the well-known[21] theorem that if

$$\int f(\rho)e^{-i\omega\rho}d\rho=F(\omega), \quad \text{(A6)}$$

then

$$\int\frac{f(\rho)}{-i\rho}e^{-i\omega\rho}d\rho=\int_{-\infty}^{\omega}F(\omega')d\omega'. \quad \text{(A7)}$$

We use the following ploy: First we take the Fourier transform of Eq. (A5) ignoring the $-i\rho$ in the denominator. Then we do the angular integral. Finally, we integrate with respect to ω' as indicated by Eqs. (A6) and (A7).

The first step immediately yields

$$\frac{4\pi g^2}{3r_0^3\epsilon q}\int_0^1 d\cos\theta\left(\frac{\sin^2\theta\cos^2\theta}{3\cos^2\theta-1}\right)\times\{\delta[\omega'+\epsilon q(3\cos^2\theta-1)]-\delta[\omega'-\epsilon q(3\cos^2\theta-1)]\}. \quad \text{(A8)}$$

Since the integrand is symmetric in $\cos\theta$, we have written the integral over half the interval and multiplied by 2.

We now perform the angular integral in Eq. (A8). Integrals of the form in Eq. (A8) are handled by the

[21] A. Erdely, F. Oberhettinger, W. Magnus, and F. G. Tricomi, *Tables of Integral Transforms* (McGraw-Hill Book Company, Inc., New York, 1954), Vol. I, Chaps. 1 and 2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

following general procedure:

$$\int_a^b G(x)\delta[F(x)]dx=\int_{F(a)}^{F(b)} G(x)\,(dF/dx)^{-1}\delta(F)dF$$

$$=\frac{G}{(dF/dx)}\bigg|_{F(x)=0} \quad \text{if } F(x)=0 \text{ at some point in the interval } a\leqq x\leqq b \tag{A9}$$

$$=0, \quad \text{if } F(x)\neq 0 \text{ at every point in the interval } a\leqq x\leqq b.$$

Using the first δ function in Eq. (A8), the correspondence with Eq. (A9) is

$$G(\cos\theta)=\frac{4\pi g^2}{3r_0^3\epsilon q}\,\frac{\sin^2\theta\cos^2\theta}{3\cos^2\theta-1}, \tag{A10}$$

$$F(\cos\theta)=\omega'+\epsilon q(3\cos^2\theta-1), \tag{A11}$$

$$\frac{dF}{d\cos\theta}=6\epsilon q\cos\theta. \tag{A12}$$

From Eq. (A11) we have $F=0=>\cos\theta=(\epsilon q-\omega'/3\epsilon q)^{1/2}$, and after some algebra,

$$\frac{G}{(dF/d\cos\theta)}\bigg|_{F=0}=-\frac{2\pi g^2}{27\sqrt{3}r_0^3\epsilon^2q^2}\left(1-\frac{\omega'}{\epsilon q}\right)^{1/2}\left(1+\frac{2\epsilon q}{\omega'}\right). \tag{A13}$$

This result is valid if $[\omega'+\epsilon q(3\cos^2\theta-1)]=0$ at some point on the interval $0\leqq\cos\theta\leqq 1$, or in other words, if $-2\epsilon q<\omega'<\epsilon q$. If ω' lies outside these limits, the result is 0.

We have completed the angular integral. The expression (A13) corresponds to $F(\omega')$ in Eq. (A7). It remains to integrate over ω'. The main point is that the integral over ω', from $-\infty$ to ω, must be considered separately for three cases:

(1) $\omega<-2\epsilon q$: The integrand is always 0 and the integral vanishes.

(2) $-2\epsilon q<\omega<\epsilon q$: The integrand is given by Eq. (A13) and the limits are $-2\epsilon q$ to ω.

(3) $\omega>\epsilon q$: The integrand is given by Eq. (A13) and the limits are $-2\epsilon q$ to ϵq.

The integration is straightforward. We quote only the results

$\omega<-2\epsilon q$: 0; (A14a)

$-2\epsilon q<\omega<\epsilon q$:

$$\frac{4\pi g^2}{27r_0^3\epsilon q}\left[-\frac{1}{3}\left(1-\frac{\omega}{\epsilon q}\right)^{1/2}\left(5+\frac{\omega}{\epsilon q}\right)\right.$$

$$\left.-\ln\left|\frac{[1-(\omega/\epsilon q)]^{1/2}-1}{[1-(\omega/\epsilon q)]^{1/2}+1}\right|+\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1}\right]; \tag{A14b}$$

$\omega>\epsilon q$:

$$\frac{4\pi g^2}{27r_0^3\epsilon q}\left[\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1}\right]. \tag{A14c}$$

These are the results for the first δ function in Eq. (A8). The computation for the second δ function is carried out analogously. The result is

$\omega<-\epsilon q$: 0; (A15a)

$-\epsilon q<\omega<2\epsilon q$:

$$-\frac{4\pi g^2}{27\sqrt{3}r_0^3\epsilon q}\left[\frac{1}{3}\left(1+\frac{\omega^{1/2}}{\epsilon q}\left(5-\frac{\omega}{\epsilon q}\right)\right)\right.$$

$$\left.+\ln\left|\frac{[1+(\omega/\epsilon q)]^{1/2}-1}{[1+(\omega/\epsilon q)]^{1/2}+1}\right|\right]; \tag{A15b}$$

$\omega>2\epsilon q$:

$$-\frac{4\pi g^2}{27\sqrt{3}r_0^3\epsilon q}\left[\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1}\right]. \tag{A15c}$$

Finally, combining the results of Eqs. (A14) and (A15), we arrive at the function given in Eq. (53) in the text.

Case d. If the $\Delta\omega$'s do not occur in pairs, we revert to Eq. (46) in the text rather than to Eq. (A1). We note that the only difference between Eq. (46) and Eq. (A1) is that in Eq. (46) there is a 1 in place of one of the exponentials in Eq. (A1). The computation is therefore essentially unaltered from *Case c*, the same basic arguments applying step for step. The result of the first step [Eq. (58)] will now have one of its δ functions replaced by $\delta(\omega')$. The angular integration for that portion of Eq. (A8) which now contains $\delta(\omega')$ offers no subtleties. The final integration over ω' merely replaces the $\delta(\omega')$ by a unit step at $\omega=0$. The result for this portion of the integral is

$\omega<0$: 0; (A16a)

$\omega>0$:

$$\frac{4\pi g^2}{27\sqrt{3}r_0^3\epsilon q}\left[\sqrt{3}+\ln\frac{\sqrt{3}-1}{\sqrt{3}+1}\right]. \tag{A16b}$$

By combining Eqs. (A15) and (A16), we arrive at the $\chi(\omega)$ given by Eq. (54) in the text.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

General Theory of Cross Relaxation. II. Higher Order Processes*

W. J. C. GRANT
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received 30 January 1964)

The theory of part I is extended to multiple spin processes. The lowest order two-spin process yields a cross-relaxation probability W_{CR} whose principal characteristics are the same as those of W_{CR} for single spin flips. Higher order two-spin processes occur with much lower probability and this probability is not sharply peaked at harmonic coincidence. Three-spin and higher multiple K-spin processes yield functions $W_{CR}(\omega)$ whose width becomes progressively larger and less dependent on the concentration, and whose magnitude depends on the concentration n as n^K. The effects of short-range forces (exchange) are easily embodied in the theory, but in the absence of such forces special attention must be paid to the dipole fields of near neighbors, which are then likely to dominate the cross-relaxation process.

I. INTRODUCTION

IN part I[1] we presented a general formalism to describe the effect of the dipolar interaction on the otherwise uncoupled spins in a lattice. We applied our considerations particularly to nonradiative single-spin transitions. In the extension of the theory to higher order processes, we presuppose the entire machinery of part I.

In generalizing the previous results to more complex processes, we follow these steps: (1) generalize the statistical procedure developed in part I, (2) discuss two-spin processes in detail, (3) consider triple and higher multiple processes, and (4) consider subsidiary problems arising from near-neighbor interactions and inhomogeneous broadening.

II. STATISTICAL PROCEDURE

The matrix element used for single-spin flips was $\sum_i M_{1i}^2$. This element corresponded to an interaction in which spin one interacts with all the other spins in such a manner that spin one flips and all the others stay put. The spin variables of M_{1i} have a form like $S_-{}^1S_z{}^i$.

Now let us consider a double flip. By a double flip we mean a process in which spin one flips down and any one of the others flips up. What is the appropriate statistical procedure and what matrix element do we embed in this procedure?

The statistics breaks up into two parts. First, we consider the interaction between spin one and some fixed spin p. The matrix element has the form $M_{1p}(\mathbf{r}_{1p})$ with spin variables typically of the form $S_-{}^1S_+{}^p$. With this one-p combination fixed, we count the possible configurations of all the remaining spins. Secondly, we sum M_{1p} over all possible interactions. When one and p interact in a double flip, their interaction with one another is evidently on a different footing from their interaction with all the other spins.

A closely related question concerns the indistinguishability of the spins. Two spins are distinguishable if any of the following conditions obtain: (1) They belong to a different species; (2) they occupy different Zeeman states; (3) they occupy different positions in the lattice. Conditions (1) and (2) are obvious. The implication of condition (3) is that we do not get a new configuration if we interchange spin q for spin p, but we do get a new configuration if we move spin p (or spin q) to a different lattice site.

Suppose that the distinguishability conditions (1) and (2) give N_1 spins of kind 1, N_2 of kind 2, and so forth, with a total of N spins. Suppose we fix the positions of A spins: A_1 of kind 1, A_2 of kind 2, and so forth. We now distribute the remaining $N-A$ spins randomly into the remaining $M-A$ lattice sites. The number of configurations is clearly

$$C(A)=\frac{(M-A)!}{(N_1-A_1)!(N_2-A_2)!\cdots(M-N)!}. \quad (1)$$

By Stirling's approximation, this becomes

$$C(A)=\frac{(M-A)^{M-A}e^{-N+A}}{(N_1-A_1)!(N_2-A_2)!(M-N)^{M-N}}. \quad (2)$$

Now

$$\lim_{M\to\infty}(M-A)^{M-A}=\lim_{M\to\infty}M^{M-A}\times(1-(A/M))^{M-A}=M^{M-A}e^{-A}, \quad (3)$$

so that (putting $n=N/M$)

$$C(A)=[M^{N-A}/(N_1-A_1)!(N_2-A_2)!\cdots]\times[e^{-N}/(1-n)^{M-N}]. \quad (4)$$

We notice that the second bracket of Eq. (4) is independent of A and will always cancel out when we compute a fractional number of configurations. For a two-spin process involving spin 1 and spin p, we have $A_1=A_2=1$, $A=2$, and the fractional number of configurations is

$$f=\frac{C(2)}{C(0)}=\frac{N_1N_2v^2}{V^N}\int d\mathbf{r}_2d\mathbf{r}_3\cdots d\mathbf{r}_{p-1}d\mathbf{r}_{p+1}\cdots d\mathbf{r}_N, \quad (5)$$

* This work was supported in part by the U. S. Army Signal Corps. under Contract DA-36-039-sc-89169.

[1] W. J. C. Grant, Phys. Rev. 134, A1554 (1964), part I, preceding paper.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where

$$v=V/M \tag{6}$$

and V has been replaced by $\int d\mathbf{r}$ in the numerator.

To obtain the transition probability, we need to introduce the transition operator and the appropriate energy constraint. The transition operator, as we have seen, is $M_{1p}^2(\mathbf{r}_{1p})$ summed over all $\mathbf{r}_{1p}$. The energy constraint is discussed in the next section. The transition probability may be symbolically written

$$W=\frac{2\pi}{\hbar}\frac{N_1N_2v^2}{V^N}\cdot\frac{1}{v}$$
$$\times\int M_{1p}^2 d\mathbf{r}_{1p}d\mathbf{r}_{12}\cdots d\mathbf{r}_{1,\,p-1}d\mathbf{r}_{1,\,p+1}\cdots d\mathbf{r}_{1N}$$
$$\times\text{energy constraint.} \tag{7}$$

We have introduced a factor $1/v$ to keep the units and dimensionality consistent in passing from the lattice sum for M_{1p}^2 to an integral. Also, the reference spin 1 has been chosen as origin of coordinates.

For a single-spin process, $C(1)/C(0)$ will have one N in front, and one more volume integral, so that W reduces to the expression 22 of part I.

The extension of this counting scheme to processes involving an arbitrary number of spins is obvious, and some examples will be given below.

III. TWO-SPIN PROCESSES

First, we formulate the energy constraint for the two-spin process. Second, we dicusss the first-order two-spin process. Third, we discuss a second-order two-spin process. Fourth, we consider the long-range process.

Energy Constraint

In a two-spin process, spin one flips from state A to state B, while another spin flips from state C to state D. The labels A, B, C, D designate the quantum numbers of the unperturbed single-particle states. The corresponding unperturbed energies are $E_A{}^0, E_B{}^0, E_C{}^0, E_D{}^0$. These energies are perturbed by the dipole interaction, and in first order we write

$$E^{\text{tot}}=E^0+E^{\text{dip}}. \tag{8}$$

E^{tot} and E^0 are numbers, but E^{dip} depends both on the specific configuration of all dipoles which the spin sees, and on the quantum state in which the spin finds itself. For a transition to be possible, we require zero change of over-all energy. The net change in Zeeman energy must be balanced by the net change in dipole energy:

$$(E_A{}^{\text{tot}}-E_B{}^{\text{tot}})+(E_C{}^{\text{tot}}-E_D{}^{\text{tot}})=0, \tag{9}$$

Equation (9) replaces the analogous condition $E_A{}^{\text{tot}}-E_B{}^{\text{tot}}=0$, which was appropriate for single flips.

To formulate the analogous energy constraint, we specify that: (1) spin 1 have energy $E_A{}^{\text{tot}}$ when in state A; (2) spin 1 have energy $E_B{}^{\text{tot}}$ when in state B; (3) spin p have energy $E_C{}^{\text{tot}}$ when in state C; (4) spin p have energy $E_D{}^{\text{tot}}$ when in state D; and (5) the constraint expressed in Eq. (9) holds. The total number of configurations for which Eq. (9) holds is obtained by integrating over all values of $E_A{}^{\text{tot}}, E_B{}^{\text{tot}}, E_C{}^{\text{tot}}, E_D{}^{\text{tot}}$. The energy constraint is then expressed as the operator

$$\int\delta[E_A{}^{\text{tot}}-(E_A{}^0+E_A{}^{\text{dip}})]\delta[E_B{}^{\text{tot}}-(E_B{}^0+E_B{}^{\text{dip}})]$$
$$\times\delta[E_C{}^{\text{tot}}-(E_C{}^0+E_C{}^{\text{dip}})]\delta[E_D{}^{\text{tot}}-(E_D{}^0+E_D{}^{\text{dip}})]$$
$$\times\delta[E_A{}^{\text{tot}}-E_B{}^{\text{tot}}+E_C{}^{\text{tot}}-E_D{}^{\text{tot}}]$$
$$\times dE_A{}^{\text{tot}}dE_B{}^{\text{tot}}dE_C{}^{\text{tot}}dE_D{}^{\text{tot}}. \tag{10}$$

Performing the integrations reduces expression (10) to

$$\delta[(E_A{}^0-E_B{}^0+E_C{}^0-E_D{}^0)$$
$$-(E_B{}^{\text{dip}}-E_A{}^{\text{dip}}+E_D{}^{\text{dip}}-E_C{}^{\text{dip}})]. \tag{11}$$

The first parenthesis in expression (11) is the total change in unperturbed energy and we denote it by $\hbar\omega$. The second parenthesis is the negative of the change in dipole energy. We put

$$\hbar\Delta\omega_{AB}=E_B{}^{\text{dip}}-E_A{}^{\text{dip}}, \tag{12a}$$
$$\hbar\Delta\omega_{CD}=E_D{}^{\text{dip}}-E_C{}^{\text{dip}}. \tag{12b}$$

We again assume that the contributions of all spins to $\Delta\omega$ are additive:

$$\Delta\omega_{AB}=\sum_{k\neq1}\Delta\omega_{AB}(\mathbf{r}_{1k}), \tag{13a}$$
$$\Delta\omega_{CD}=\sum_{k\neq p}\Delta\omega_{CD}(\mathbf{r}_{pk}). \tag{13b}$$

Expression (11) now becomes

$$\delta(\omega-[\sum_{k\neq1}\Delta\omega_{AB}(\mathbf{r}_{1k})+\sum_{k\neq p}\Delta\omega_{CD}(\mathbf{r}_{pk})]). \tag{14}$$

Finally, replacing the δ function by its Fourier transform, we have for the energy constraint:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\{-i\rho[\omega-\sum_{k\neq1}\Delta\omega_{AB}(\mathbf{r}_{1k})-\sum_{k\neq p}\Delta\omega_{CD}(\mathbf{r}_{pk})]\}d\rho$$
$$\equiv\frac{1}{2\pi}\int_{-\infty}^{\infty}(\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1p})+\Delta\omega_{CD}(\mathbf{r}_{p1})]\}$$
$$\times\prod_{k\neq1,p}\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1k})+\Delta\omega_{CD}(\mathbf{r}_{pk})]\})e^{-i\rho\omega}d\rho. \tag{15}$$

First-Order Process

We may now write the transition probability in the form of Eq. (7) using the energy constraint we have

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

just derived. We put $N_i/V=n_i/v$. We then have

$$W(\omega)=\frac{2\pi}{\hbar}\int_{-\infty}^{\infty}\frac{1}{2\pi}\Big[n_1n_2\cdot\frac{1}{v}\int M_{1p}{}^2\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1p})+\Delta\omega_{CD}(\mathbf{r}_{p1})]\}d\mathbf{r}_{1p}\Big(\frac{1}{V}\int\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1k})+\Delta\omega_{CD}(\mathbf{r}_{pk})]\}d\mathbf{r}_{1k}\Big)^{N-2}\Big]e^{-i\omega\rho}d\rho. \quad (16)$$

This expression is very similar to Eq. (24) of part I, with the integral containing $M_{1p}{}^2$ corresponding to χ and the round bracket raised to the $(N-2)$ power corresponding to Φ. The difference is that χ and Φ cannot now be evaluated separately. They are coupled through the dependence of Φ on $\mathbf{r}_{1p}$. This becomes more apparent if we write $\mathbf{r}_{pk}=\mathbf{r}_{1k}-\mathbf{r}_{1p}$. As we integrate over $\mathbf{r}_{1p}$ to evaluate χ, we should simultaneously have to integrate out the $\mathbf{r}_{1p}$ dependence in Φ. This calculation would be impossible.

This impasse can be overcome through the following consideration: Let us consider the limiting situations, $\mathbf{r}_{1p}\rightarrow 0$ and $\mathbf{r}_{1p}\rightarrow\infty$. In the first case, the $\mathbf{r}_{1p}$ dependence of Φ is eliminated and Φ has exactly the form we discussed in part I. It will essentially describe a quasiresonance line whose parameters are the algebraic sum of the parameters for the resonance lines $A\leftrightarrow B$ and $C\leftrightarrow D$. In the second case, we note that $\Delta\omega(\mathbf{r})$ depends on r as r^{-3}. Within any arbitrarily large neighborhood of spin 1, $\mathbf{r}_{pk}$ will be infinite, $\Delta\omega(\mathbf{r}_{pk})$ will be zero, and $\exp[i\rho\Delta\omega(\mathbf{r}_{pk})]$ will equal one. The converse is true within any arbitrarily large neighborhood of spin p. Consequently, each of the $(N-2)$ integrals in Φ breaks up into the sum of two integrals.

$$\lim\Phi_{r_{1p}\rightarrow\infty}=\Big(\frac{1}{V}\int\{\exp[i\rho\Delta\omega_{AB}(\mathbf{r})]+\exp[i\rho\Delta\omega_{CD}(\mathbf{r})]\}dr\Big)^{N-2}. \quad (17)$$

It is easily shown that Φ now corresponds to the convolution of the two resonance lines $A\leftrightarrow B$ and $C\leftrightarrow D$.

In either limit, Φ is functionally identical to an absorption line and has a width roughly twice of an absorption line. For most practical applications, this is sufficient information. If more detail is necessary one must split the calculation into two parts corresponding to $\mathbf{r}_{1p}$ very small and $\mathbf{r}_{1p}$ very large. For $\mathbf{r}_{1p}$ small it is always possible to calculate exactly the diagonal dipole elements which determine $\Delta\omega$. For $\mathbf{r}_{1p}$ large it is always possible to convolute known absorption lines.

What is the physical meaning of the preceding arguments? In the first place it is no longer exactly true that χ and Φ represent, respectively, the rf power spectrum of the dipole interaction and the level broadening. Such a separation was possible for the single-spin theory of part I because of our assumption that the interactions of spin 1 were independent. Now, however, the interactions of spin 1 are not completely independent; they are correlated with the interactions of spin p. From our present point of view, we still assign to χ a meaning as power spectrum, and take care of the correlations in Φ. In line shape language, $\Phi(\rho)$ is the coherence relaxation function and its width corresponds to T_2. When two spins are very near one another, they are very strongly correlated. With respect to the rest of the system, they behave as a coupled unit with a coherence time equal to that of a single spin. When two spins are very far apart, on the other hand, they are uncorrelated, and their joint coherence time is given by

$$1/T_2(1,2)=[1/T_2(1)]+[1/T_2(2)],$$

or, if $T_2(1)=T_2(2)$,

$$T_2(1,2)=\tfrac{1}{2}T_2(1)=\tfrac{1}{2}T_2(2).$$

The dependence of Φ on r_{1p} is thus interpretable as the dependence of the *pair* coherence time on the spin separation.

We may then write for the double-flip probability

$$W(\omega)=\frac{2\pi}{\hbar}\frac{n_1n_2}{v}\int_{-\infty}^{\infty}\chi(\rho)\Phi(\rho)e^{-i\rho\omega}d\rho, \quad (18)$$

where

$$\chi(\rho)=\frac{1}{2\pi}\int M_{1p}{}^2e^{i\rho\Delta\omega_\chi(\mathbf{r})}d\mathbf{r}, \quad (19)$$

$$\Delta\omega_\chi(\mathbf{r})=\Delta\omega_{AB}(\mathbf{r}_{1p})+\Delta\omega_{CD}(\mathbf{r}_{p1}). \quad (20)$$

We now discuss briefly the more precise functional form of M_{1p} and $\Delta\omega$ appearing in Eqs. (18) to (20). The matrix element M_{1p} is, of course,

$$\langle A(1)C(p)|\mathcal{H}^{\mathrm{dip}}|B(1)D(p)\rangle.$$

In the absence of a crystal field, we would have

$$M_{1p}=g(3\cos^2\theta_{1p}-1)/r_{1p}{}^3, \quad \Delta m_{tot}=0, \quad (21a)$$

$$M_{1p}=g\sin^2\theta_{1p}/r_{1p}{}^3, \quad \Delta m_{tot}=\pm 2. \quad (21b)$$

The quantity g is the value of the appropriate spin matrix element. Like the matrix element for single flips, M_{1p} can be expressed even in the most general case as the product of an angular function and r^{-3}.

$$M_{1p}=m(\theta_{1p})/r_{1p}{}^3. \quad (22)$$

The $\Delta\omega$'s have the same form as in the single-spin process. In the absence of a crystal field

$$\Delta\omega_{AB}(\mathbf{r}_{1k})=q_{AB}(3\cos^2\theta_{1k}-1)/r_{1k}{}^3, \quad (23)$$

with an analogous expression for $\Delta\omega_{CD}$. The q_{AB} is the difference of the values assumed by the spin variables in the matrix elements $\mathcal{H}_{AA}{}^{\mathrm{dip}}$ and $\mathcal{H}_{BB}{}^{\mathrm{dip}}$. In the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

general case, $\Delta\omega$ too can still be expressed as the product of an angular function and r^{-3}:

$$\Delta\omega(\mathbf{r}_{1k}) = f(\theta_{1k})/r_{1k}^{-3}. \tag{24}$$

We see then that the χ of Eqs. (18) and (19) is identical in form to the single-spin χ of Eqs. (25) and (38) in part I. The functions $m(\theta)$ and $f(\theta)$ are still sums of Legendre polynomials of order two.

The entire discussion χ developed in part I applies without essential change to the two-spin case. Thus, we may write

$$\begin{aligned}\chi(\omega) &= \frac{8\pi\langle m^2\rangle}{3r_0^3}\frac{1}{2L\epsilon}\ln|L\epsilon/\omega|\,, \quad |\omega|<L\epsilon \\ &= 0, \qquad\qquad\qquad\quad |\omega|>L\epsilon.\end{aligned} \tag{25}$$

Here, $\langle m^2\rangle$ is the average value of $m^2(\theta)$ over the sphere, L is the maximum value of $f(\theta)$, and $\epsilon = r_0^{-3}$. Unlike the single flip, the double flip is not characterized by the occurrence of $\Delta\omega$ and $-\Delta\omega$ in pairs. The resulting χ is asymmetric and this asymmetry is well approximated by the addition of an additive constant on one side of the origin. As mentioned in part I, such asymmetry manifests itself primarily in the wings.

The discussion of $W(\omega)$ developed in part I also applies, with two changes: (1) There is now a coefficient of n^2 instead of n, so that the transition probability depends on the square of the concentration. (2) When calculating the central peak of W, we must use a Φ that is wider than its counterpart in the single-spin transition. The theory developed in part I regarding the shape of $W(\omega)$ and the concentration dependence of this shape applies to the double flip without further modification.

Second-Order Process

So far, we have treated the two-spin process in terms of first-order theory. It may happen, however, that the "resonant" condition is met, i.e., $\Delta E_{AB}{}^0 + \Delta E_{CD}{}^0 = 0$, or $\omega = 0$, but at the same time the matrix element M_{1p} vanishes. Such is the case for ruby at a number of magnetic field values when the field is parallel to the c axis. To see what effect is expected under these circumstances, we must go to second-order theory.

The second-order matrix element is given by

$$M_{1p}{}' = \sum_{X\neq B}\sum_{Y\neq D}\frac{\langle A(1)C(p)|\mathcal{H}^{\rm dip}|X(1)Y(p)\rangle\langle X(1)Y(p)|\mathcal{H}^{\rm dip}|B(1)D(p)\rangle}{E_{XY}-E_{BD}}, \tag{26}$$

where X and Y are the intermediate states for spins 1 and p. Each of the first-order elements connecting the intermediate states to either the initial or final state is of the form of M_{1p} given in Eqs. (21) and (22). Equation (26) may be abbreviated in the notation of Eq. (22),

$$M_{1p}{}' = \sum_{XY}\frac{[m_1(\theta_{1p})m_2(\theta_{1p})/r_{1p}{}^6]}{\Delta E_{XY}}. \tag{27}$$

In other words, $M_{1p}{}'$ breaks up into products of an angular function, r^{-6}, and an energy denominator. The energy denominator is

$$\begin{aligned}1/\Delta E = 1/(\Delta E^0 + \Delta E^{\rm dip}) &= (1/\Delta E^0) \\ &\times[1-(\Delta E^{\rm dip}/\Delta E^0)+(\Delta E^{\rm dip}/\Delta E_0)^2\cdots]\,.\end{aligned} \tag{28}$$

Since, for the bulk of the transitions, $\Delta E^{\rm dip}$ is small compared to ΔE^0, we approximate the energy denominator by $1/\Delta E^0$.

The number of terms in the XY sum is generally small. The reason for this is that the vanishing of the first-order matrix element is generally due to some symmetry condition which causes many of the other matrix elements to vanish as well. For instance, in the case of ruby at 0° orientation, for $g\beta H = 4D$, the transition $(\frac{3}{2},-\frac{3}{2})\leftrightarrow(-\frac{1}{2},-\frac{1}{2})$ is linked by two intermediate states. Both terms in the sum over intermediate states are proportional to $P_{20}P_{21}r^{-6}$. The coefficients of the coordinate functions are 0.4 and -2.4 so that their sum is -2. We cite this example to illustrate that the order of magnitude of the sum is the same as the order of magnitude of each term.

We now proceed exactly as in the first-order theory. The integral for χ' is now

$$\chi'(\rho) = \frac{1}{2\pi}\cdot\frac{1}{(\Delta E)^2}\int_{r_0}^{\infty}\frac{m_1{}^2(\theta)m_2{}^2(\theta)}{r^{12}}e^{i\rho f(\theta)/r^3}d\mathbf{r}. \tag{29}$$

We identify this integral with the integral $I(\rho,4)$ of Eq. (38) in part I. Then, according to Eq. (43) of part I, the integration over r (and φ) yields

$$\chi'(\rho) = \left(\frac{1}{\Delta E}\right)^2\cdot\int_{-1}^{1}d(\cos\theta)\left(\frac{m_1{}^2(\theta)m_2{}^2(\theta)}{3r_0{}^9}\right)\times\frac{\gamma[3,-i\rho\epsilon f(\theta)]}{[-i\rho\epsilon f(\theta)]^3}, \tag{30}$$

where γ represents the incomplete γ function. We approximate $f(\theta)$, as before, by a function which assumes values $-L<f<L$ as $\cos\theta$ assumes values $-1<\cos\theta<1$, and we replace $m_1{}^2(\theta)m_2{}^2(\theta)$ by $\langle m_1{}^2m_2{}^2\rangle$. With $\epsilon = r_0^{-3}$, we then have

$$\chi'(\rho) = \frac{1}{\epsilon}\cdot\left(\frac{1}{\Delta E}\right)^2\frac{\langle m_1{}^2m_2{}^2\rangle\epsilon^4}{3L}\int_{-L}^{L}df\frac{\gamma(3,-i\rho\epsilon f)}{(-i\rho\epsilon f)^3}. \tag{31}$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Through the relation[2]

$$\int \frac{\gamma(a,x)}{x^a}dx = -\frac{\gamma(a-1,x)}{x^{a-1}}, \tag{32}$$

we obtain

$$\chi'(\rho) = \frac{1}{\epsilon}\cdot\left(\frac{1}{\Delta E}\right)^2 \cdot \frac{\langle m_1^2 m_2^2\rangle \epsilon^4}{3}\cdot\left[\frac{\gamma(2,-x)-\gamma(2,x)}{x^3}\right], \tag{33}$$

where $x=-i\rho\epsilon L$. By means of the recursion formula[2]

$$\gamma(a+1,x) = a\gamma(a,x) - x^a e^{-x} \tag{34}$$

and the relation

$$\gamma(1,x) = 1-e^{-x}, \tag{35}$$

Eq. (33) may be reduced to

$$\chi'(\rho) = \frac{1}{\epsilon}\left(\frac{1}{\Delta E}\right)^2 \frac{\langle m_1^2 m_2^2\rangle \epsilon^4}{3} \times\left[e^{-x}\left(\frac{1}{x^3}+\frac{1}{x^2}\right) - e^{x}\left(\frac{1}{x^3}-\frac{1}{x^2}\right)\right]. \tag{36}$$

We have now explicitly evaluated $\chi'(\rho)$.

To obtain $\chi'(\omega)$ we need

$$\chi'(\omega) = \int_{-\infty}^{\infty} \chi'(\rho) e^{-i\omega\rho} d\rho. \tag{37}$$

This Fourier transform is obtained almost trivially from expression (36) by taking successive integrals of δ functions located at $\omega = \pm\epsilon L$. The Fourier transform of expression (36) is

$$\chi'(\omega) = \frac{2\pi\langle m_1^2 m_2^2\rangle \epsilon^2}{9 r_0^3 (\Delta E)^2}\frac{3}{4L\epsilon}\left[1-\left(\frac{\omega}{L\epsilon}\right)^2\right], \quad |\omega| < L\epsilon$$
$$= 0, \quad |\omega| > L\epsilon. \tag{38}$$

$\chi'(\omega)$ turns out to be a parabola. It has a half-width of $0.7L\epsilon$. For $L=2q$, as is appropriate for ruby at 0° orientation, this half-width corresponds to 1.4 times the effective nearest-neighbor dipole energy. This width is much larger than the width of any reasonable $\Phi(\omega)$ at low concentrations. It follows that $W'(\omega)$, obtained by convoluting Φ and χ', will have essentially the same shape as $\chi'(\omega)$.

The coefficient in front of the second-order W' is down from that of the first-order W by $\frac{1}{6}(\bar{m}\epsilon/\Delta E_0)^2$. Since $\bar{m}\epsilon$ is approximately the dipole energy and ΔE_0 is approximately the Zeeman energy, this factor is typically of order 10^{-2}. This is not all, however. The second-order $W'(\omega)$ shows no sharp resonance. If we consider a $\Phi(\omega)$ of width one-tenth the nearest-neighbor dipole energy, the sharp resonance of the first-order W decreases the ratio of $W'(0)/W(0)$ by almost another order of magnitude. Far away from the resonant condition, however, W' eventually may become comparable to W.

In summary, the second-order process gives rise to an effect which is relatively small, as expected, and which is quite spread out in its frequency dependence. One would not expect this process to be directly observable, but it does contribute, to some extent, to "general" cross-relaxation, i.e., to relaxation which appears more or less independent of any resonance condition between levels.

Long-Range Process

Finally, we consider still a third type of two-spin process. In this process, spin 1 and spin p both flip, but independently of one another. It represents, in a sense, the joint probability of two single spin flips. Because the two spins operate independently, their separation becomes irrelevant. Our theory gives a reasonable account of the nonexistence of such a process.

The matrix element governing this process is

$$M_{1p}'' = \sum_{XY}\left\{\frac{\langle A(1)X(i)|\mathcal{H}^{\rm dip}|B(1)X(i)\rangle\langle C(p)Y(j)|\mathcal{H}^{\rm dip}|D(p)Y(j)\rangle}{E_C-E_D} + \frac{\langle C(p)Y(j)|\mathcal{H}^{\rm dip}|D(p)Y(j)\rangle\langle A(1)X(i)|\mathcal{H}^{\rm dip}|B(1)X(i)\rangle}{E_A-E_B}\right\}. \tag{39}$$

The two portions of this matrix element correspond to the two possibilities of either spin 1 or spin p flipping first, thus giving rise to different intermediate states. Since our method always envisions *exact* conservation of energy in any one process, one energy denominator is the *exact* negative of the other. By abbreviating the notation in an obvious manner, we have

$$M_{1p}'' = \frac{1}{\Delta E}\sum_{XY}(M_{1i}M_{pj} - M_{pj}M_{1i}). \tag{40}$$

If $i\neq j$, the spins referred to in M_{1i} and M_{pj} are all different, M_{1i} and M_{pj} commute, and M_{1p}'' vanishes identically. If $i=j$, M_{1p}'' does not vanish. However, the extra index "uses up" one more volume integral in our configuration counting [see Eqs. (7) and (16)], and as a result, there is an extra factor of $1/V$ left over. This leads to a transition probability which is down by a factor of N, where N is the number of spins in the crystal. Physically, spin 1 and spin p cannot interact at all without any interaction mechanism. They can interact via some common third spin i, but since they act independently, the chances of their "picking" out

[2] A. Erdely, F. Oberhettinger, W. Magnus, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1948), Vol. 2.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the same spin i from the N available spins are down by a factor of N.

This completes our treatment of the two-spin process.

IV. HIGHER ORDER PROCESSES

Triple and higher multiple spin flips are handled in a fashion quite analogous to double flips. The principal difference is that dipole operators can never connect more than two spins at a time; for this reason, multiple-spin processes must be treated in higher order perturbation theory.

We consider a triple flip in which spin 1 goes from state A to state B, spin p from C to D, spin q from E to F. If the selection rule $\Delta m=\pm 1$ holds for each spin, the possible processes are that either spin 1 or spin 2 or spin 3 flips first with the remaining two spins flipping second, or the converses of these three possibilities. There are thus six intermediate states, with a given process and its converse having energy denominators that are exact negatives of one another. These processes have been discussed in their detailed application to LiF by Pershan.[3] In the more general case, when m is not a good quantum number, there is no selection rule. The number of processes can become very large and the matrix elements very complicated. Nevertheless, a typical term, together with its converse, can be represented by

$$M_{1pq}=(M_{1p}M_{1q}-M_{1q}M_{1p})/\Delta E\,. \tag{41}$$

M_{1pq} does not vanish since M_{1p} and M_{1q} do not commute. The coordinate dependence of M_{1p} and M_{1q} is defined in Eqs. (21) and (22). We therefore have

$$M_{1pq}=\frac{1}{\Delta E}\frac{m_1(\theta_{1p})m_2(\theta_{1q})}{r_{1p}{}^3r_{1q}{}^3}\,. \tag{42}$$

When following our previous reasoning, the operator for a triple flip becomes

$$M^2=\left(\frac{1}{v\Delta E}\right)^2\int\frac{m_1{}^2(\theta_{1p})m_2{}^2(\theta_{1q})}{r_{1p}{}^6r_{1q}{}^6}d\mathbf{r}_{1p}d\mathbf{r}_{1q}\,. \tag{43}$$

In our configuration counting, we must include the fact that we have already assigned three spins to three "boxes" and thus, the analog of Eq. (5) is

$$f=\frac{C(3)}{C(0)}=\frac{N_1N_2N_3v^3}{V^N}\int d\mathbf{r}_{12}\cdots d\mathbf{r}_{1,\,p-1}\times d\mathbf{r}_{1,\,p+1}\cdots d\mathbf{r}_{1,\,q-1}d\mathbf{r}_{1,\,q+1}\cdots d\mathbf{r}_{1N}\,. \tag{44}$$

In setting up the energy constraint, we require that $\Delta E^0+\Delta E^{\rm dip}=0$ when all three spins are considered. In analogy to expression (15), this constraint turns out to be

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1p})+\Delta\omega_{AB}(\mathbf{r}_{1q})+\Delta\omega_{CD}(\mathbf{r}_{p1})+\Delta\omega_{CD}(\mathbf{r}_{pq})+\Delta\omega_{FG}(\mathbf{r}_{q1})+\Delta\omega_{FG}(\mathbf{r}_{qp})]\}\times\prod_{k\neq 1,\,p,\,q}\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1k})+\Delta\omega_{CD}(\mathbf{r}_{pk})+\Delta\omega_{FG}(\mathbf{r}_{qk})]\}e^{-i\rho\omega}d\rho\,. \tag{45}$$

Finally, defining $\Delta\omega_\chi$ in analogous fashion as in Eq. (20) we have

$$W(\omega)=\frac{2\pi}{\hbar}\frac{n_1n_2n_3}{v^2}\int_{-\infty}^{\infty}\chi(\rho)\Phi(\rho)e^{-i\rho\omega}d\rho\,, \tag{46a}$$

$$\chi(\rho)=\frac{1}{2\pi}\cdot\left(\frac{1}{\Delta E}\right)^2\int\left[\frac{m_1{}^2}{r_{1p}{}^6}e^{i\rho\Delta\omega_\chi(\mathbf{r}_{1p})}\right]\times\left[\frac{m_2{}^2}{r_{1p}{}^6}e^{i\rho\Delta\omega_\chi(\mathbf{r}_{1q})}\right]e^{i\rho\Delta\omega_\chi(\mathbf{r}_{pq})}d\mathbf{r}_{1p}d\mathbf{r}_{1q}\,, \tag{46b}$$

$$\Phi(\rho)=\left(\frac{1}{V}\int\exp\{i\rho[\Delta\omega_{AB}(\mathbf{r}_{1k})+\Delta\omega_{CD}(\mathbf{r}_{pk})+\Delta\omega_{FG}(\mathbf{r}_{qk})]\}d\mathbf{r}_{1k}\right)^{N-3}. \tag{46c}$$

Once again, χ and Φ are not really separable, but our previous discussion of this point applies without change.

A new difficulty arises, however, because χ itself is now given by two integrals which are not separable. They are not separable because of the appearance of $\mathbf{r}_{pq}$ in the third exponential. To test the effect of this coupling term, we have computed $\chi(\omega)$ numerically, both with and without the coupling term. We used P_{21} for $m(\theta)$ and P_{20} for $\Delta\omega_\chi$. We find that the two χ's generated in this way have virtually the same shape, except that the χ without the coupling term is narrower by about a factor of 1.5. It appears to be still practicable, therefore, to calculate χ on the basis of separable integrals with the proviso that one expects the correct function to be qualitatively the same but somewhat broader.

Each of the separate integrals gives the power spectrum of the projection of the dipole operator governing the interaction between spin 1 and spin p and between spin 1 and spin q, respectively. If these two interactions were completely independent, there would be no coupling term, and the joint power spectrum would simply be the convolution of the individual power spectra. But the transitions do not occur independently; hence, a coupling term. In picturesque language, the relative squiggles of spins 1 and p are modulated by the squiggles of spin q, and the spectrum is broadened.

[3] P. S. Pershan, Phys. Rev. 117, 109 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

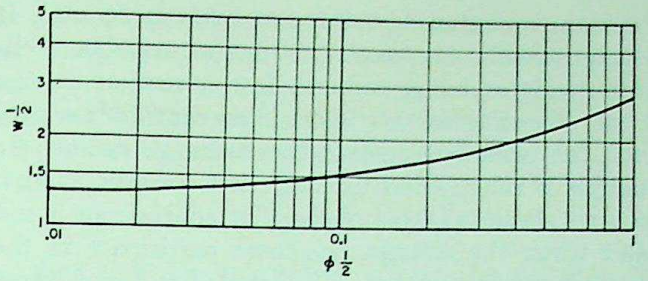


FIG. 1. Three-spin functions: W half-width versus Φ half-width in units of χ half-width.

The calculation of the triple flip $W(\omega)$ again resolves itself into the convolution of the χ and Φ functions of part I. Only now we need two χ's with their associated Φ's instead of one. We note that $W(\omega)$ will have a factor proportional to the cube of the concentration and to the square of the ratio of the dipole energy to the Zeeman energy.

Four-spin and higher order processes require no new physical or computational considerations. We note that a matrix element of the form $M_{1p}M_{qr}$ gives zero contribution. The two factors commute and the element will be identically cancelled by its converse. This is important since there is no implicit requirement in such a matrix element that spins 1 and p be near spins q and r.

When K spins flip, we will therefore require $(K-1)$ 'th-order perturbation theory, and the matrix element will be a sum of terms each of which contains $(K-1)$ products. Our counting scheme will generate a concentration dependence of order n^K, and the $(K-1)$ products of dipole operators will generate a factor of (dipole energy/Zeeman energy)$^{2(K-2)}$. The functional form can be estimated by convoluting $(K-1)$ χ functions and their associated Φ's. Owing to this increasing number of convolutions, it is clear that higher order processes will show a less and less marked resonance effect, and that their $W(\omega)$ will become more and more smeared over ω space. When this consideration is coupled with the step n dependence of these processes, we have at least a partial explanation of why the "general" cross-relaxation increases so markedly at higher concentrations.

The concentration dependence of the *shape* of $W(\omega)$ enters exclusively through the Φ functions, as we have seen in part I. We showed that $W(\omega)$ for a one-spin process has a width which varies as the square root of the Φ width. If the resonance line is homogeneous, the width of this $W(\omega)$ will vary as $n^{1/2}$. The same argument applies without change to the two-spin process. For higher multiple processes, however, the concentration dependence of the width of $W(\omega)$ is much weaker. In Fig. 1, we show the half-width of W as a function of the half-width of Φ, for a three-spin process. The curve is virtually flat until the Φ width is about one-tenth the χ width. The slope of the curve then increases gradually and reaches the value $\frac{1}{2}$ only when Φ and χ have equal widths. The concentration dependence of the width of $W(\omega)$ rapidly disappears as the number of spins involved increases. The reason for this is that the successive convolutions rapidly smear out the sharp central peak which alone depends critically on the width of Φ.

V. EXCHANGE

We have so far ignored certain considerations which do not alter the general structure of our calculation, but which nonetheless may have a pronounced effect on actual results.

An exchange interaction of the form $J(\mathbf{S}_1\cdot\mathbf{S}_2)$ frequently occurs between neighboring transition metal ions. If such an interaction is present, we distinguish two cases: (1) When the exchange energy is of the same order as the dipole energy, we treat these two interactions in parallel fashion. We still consider the unperturbed energy of the ion as an eigenvalue of a single-ion state. We consider the diagonal elements of $J(\mathbf{S}_1\cdot\mathbf{S}_2)$ as contributing to $\Delta\omega$ and we consider the off-diagonal elements as contributing to the transition matrix elements. The integral defining χ [Eq. (25), part I] will, of course, be split into two parts. The first part sums over those sites that are exchange coupled, the second over the rest of space. (2) When the exchange energy becomes comparable to the Zeeman or crystal field energy, the unperturbed states can no longer be considered as single-ion states. If J is large enough, E^0 is given by the diagonal elements in a representation in which $J(\mathbf{S}_1\cdot\mathbf{S}_2)$ is diagonal. In any case, since the energy separations between pair states are now, in general, quite different from the energy separations between single-particle states, such exchange coupled pairs do not, in general, contribute to the resonant cross-relaxation process. In terms of computational procedure, this means that the integral defining χ starts at a radius where exchange is no longer large.

This prescription is somewhat oversimplified for two reasons. First, it may happen accidentally that the separation between some of the pair levels does, in fact, coincide closely with the separation between some of the single-particle levels. Secondly, as we have seen, the frequency width of higher order processes becomes large. At the same time, the number of pair levels is $(2S+1)$ times the number of single-particle levels. As the requirement of harmonic coincidence is made less and less stringent, it becomes more and more probable that some of the level separations in the larger pair manifold will meet an approximate harmonic condition. One expects that exchange coupled pairs would, in general, enhance the probability of higher order processes, particularly under conditions where the single-particle levels alone do not exactly meet the harmonic condition. Such participation of exchange-coupled pairs in the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

cross-relaxation process has been discussed by Bloembergen and Pershan[4] and by Gill and Elliott.[5]

VI. THE NEAR-NEIGHBOR SPECTRUM

From the very beginning, we have replaced lattice sums by volume integrals, using the volume per lattice site v to keep the dimensionality consistent. Clearly, such a substitution is valid when the density of lattice sites per unit solid angle is large. For the calculation of Φ this procedure will introduce a negligible error, except possibly in certain concentrated systems. The principal reason for this is that the dominant contribution to Φ comes from distant lattice sites. For the calculation of χ, the equivalence of the discrete and "continuous" lattice is of much more restricted validity, because the dominant contribution comes from close-lying sites. It is valid when a substantial number of neighbor shells are strongly exchange coupled. In this case, as we saw in the preceding section, there is no contribution to χ until the sites fall outside the exchange radius. An example is ruby, where exchange is large over at least the first seven neighbor shells. The equivalence is not valid when the nearest neighbors participate in the cross-relaxation process. The contribution of these neighbors will dominate over the contribution of the rest of the lattice because of the r^{-6} dependence of the transition operator. The integral for χ is then replaced by a sum.

$$\chi(\rho)=\frac{1}{2\pi}\sum_j \frac{m_j^2(\theta_j)}{r_j^6}e^{i\Delta\omega_j\rho}, \tag{47}$$

where the sum need extend only as far as the first few, sometimes only the first, neighbor shells. Then $\chi(\omega)$ will be given by a series of δ functions

$$\chi(\omega)=\sum_j \frac{m_j^2(\theta_j)}{r_j^6}\delta(\omega-\Delta\omega_j)\,. \tag{48}$$

Finally, $W(\omega)$ will be given by a sum of resonance curves, each displaced from the origin by $\Delta\omega_j$ and weighted by the corresponding transition operator:

$$W(\omega)\sim\sum_j \frac{m_j^2(\theta_j)}{r_j^6}\Phi(\omega-\Delta\omega_j)\,. \tag{49}$$

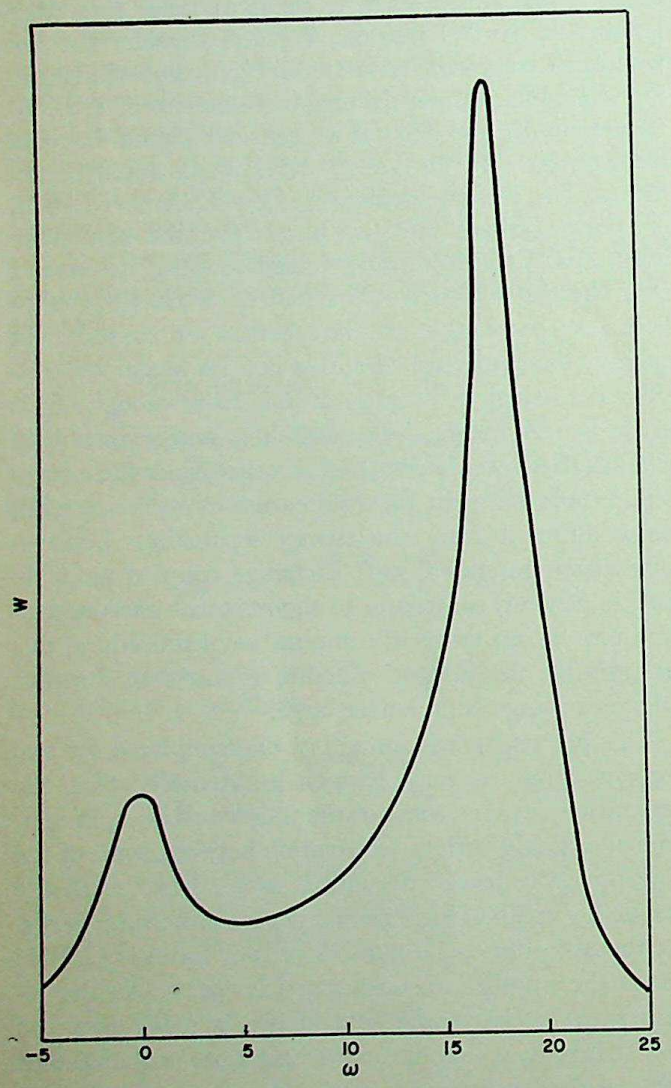


FIG. 2. $W(\omega)$ for simple cubic lattice, [001] direction, $M=\sin\theta\cos\theta$ (arbitrary units).

In systems that are very dilute and have no inhomogeneous broadening, the various Φ's may be resolvable. To illustrate, we show, in Fig. 2, $W(\omega)$ for a simple cubic lattice with the magnetic field in the [001] direction and with $\Delta\omega_j=(3\cos^2\theta_j-1)/r_j^3$, $m_j=\sin\theta_j\cos\theta_j$. The width of Φ has been taken as 1/50 the nearest-neighbor dipole energy. The peak at 18 is due to the (0,1,1) and equivalent sites. The six nearest neighbors are not effective because their m_j vanishes. If we take $m_j=(3\cos^2\theta-1)$, then the neighbors at $(0, 0, \pm1)$ will produce an effect 32 times greater than the (0,1,1) neighbors and 11 times as far away from the origin. The shape of W is quite sensitive to the angular functions involved. This means it is quite sensitive to the particular matrix element for a given process and to the direction of the magnetic field.

In more concentrated systems or in systems with substantial inhomogeneous broadening, Φ will be too broad to resolve subsidiary peaks such as shown in Fig. 2. Nevertheless, the shift of the peak away from $\omega=0$, i.e., away from the exact resonance condition, will still occur.

The shift of the peak cross-relaxation probability away from the harmonic condition becomes intuitively obvious if one considers that the matching of level separations must include the diagonal dipole energy as well as the Zeeman energy. Thus, the peak in cross re-

[4] N. Bloembergen and P. S. Pershan, in *Advances in Quantum Electronics*, edited by J. R. Singer (Columbia University Press, New York, 1961), p. 373.

[5] J. C. Gill and R. J. Elliott, in *Advances in Quantum Electronics*, edited by J. R. Singer (Columbia University Press, New York, 1961), p. 399.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

laxation still occurs where the total energies are exactly matched, but this point will be somewhat displaced from the point where the Zeeman energies alone are exactly matched. The amount of displacement depends rather critically on the specific operating conditions.

VII. INHOMOGENEOUS BROADENING

Inhomogeneous broadening affects the distribution of energy levels about their unperturbed value. In our formalism this distribution manifests itself in the function Φ. The question whether Φ should include the inhomogeneous broadening or not is isomorphic to the question whether the spin packets within an inhomogeneous line cross-relax together or separately. For limiting cases one can give an unequivocal answer. If the spin-diffusion time within the inhomogeneous line is much shorter than the cross-relaxation time for one packet, then the whole line relaxes together. If the spin-diffusion time is much longer than the cross-relaxation time, then the packets relax separately. To put this in terms of numbers, we take $T_2=10^{-8}$ sec and $T_{12}=10^{-3}$ sec. For T_{diff} we use Bloembergen's rule of thumb[6]

$$T_{\text{diff}}=T_2\times(T_2/T_2^*)^3, \tag{50}$$

where $1/T_2^*$ is the total width of the inhomogeneous line. We then obtain the condition that the homogeneous width of the line must be much greater than 2% of the inhomogeneous width, if the entire line is to relax as a whole. If this condition is met, Φ should therefore include the inhomogeneous broadening. If this condition is not met, the situation becomes complicated. Relaxation takes place between all the spin packets in one inhomogeneously broadened level and all the spin packets in the other inhomogeneously broadened level. At the same time spin diffusion takes place within each of the levels.

[6] N. Bloembergen, S. Shapiro, P. S. Pershan, and J. O. Artman, Phys. Rev. 114, 445 (1959).

VIII. SUMMARY

We have extended the basic cross-relaxation theory of part I to relaxation processes of all orders. Among other results we have derived the following: (1) For a K-spin process, the magnitude of the transition probability depends on concentration as n^K. (2) The frequency width of the transition probability is proportional to the square root of the Φ width for one- and two-spin processes. For three-spin processes, the width of W depends on the width of Φ only when Φ and χ have comparable widths, and even then the dependence is relatively feeble. For higher multiple processes, the dependence of the W width on the Φ width is negligible. Instead, the width of the K-spin W is roughly proportional to K. (3) In the absence of strong exchange interactions, the nearest neighbors dominate the cross-relaxation process. In this case, W is obtained from the convolution of Φ functions with a discrete near-neighbor spectrum, and its peak may be displaced from the harmonic point by an amount corresponding to the nearest-neighbor dipole energy.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

General Theory of Cross Relaxation. III. Application to Experiment*

W. J. C. GRANT
Bell Telephone Laboratories, Murray Hill, New Jersey

(Received 30 January 1964)

The theory developed in parts I and II is applied to the classical experiments of Mims and McGee, and Pershan. The predictions of the theory regarding the dependence of W_{CR} on the energy imbalance $\hbar\omega$ on concentration and on the exchange radius, as well as the prediction of the actual magnitude of W_{CR}, are confirmed in the Mims and McGee experiment on ruby. The crucial significance of near-neighbor dipole interactions, in particular, the effect of the associated power spectrum, which is quite broad and quite sensitive to crystal direction, is illustrated by application of the theory to Pershan's experiments on LiF.

INTRODUCTION

IN the preceding papers,[1] we have developed methods for calculating from first principles the cross-relaxation probability W_{CR} for processes of arbitrary complexity. The theory makes specific quantitative predictions about the dependence of W_{CR} on the number of spins participating in the process, on the energy imbalance involved, on the spin concentration, and on the existence of short-range forces. Our purpose now is to test the theory in its application to classical experiments in the field. We shall consider the experiments of Mims and McGee[2] on ruby and the experiments of Pershan[3] on LiF.

Transition probabilities are related to observable quantities through rate equations. The transient solutions of rate equations containing both cross-relaxation and spin-lattice terms have been discussed at length in a previous paper,[4] both in general and with particular reference to the experiments we shall discuss. We shall avail ourselves here of those results.

I. THE MIMS EXPERIMENT ON RUBY

We denote the four ground-state spin levels of Cr^{3+} in ruby by 1, 2, 3, 4 in order of decreasing energy. As discussed in detail in Mims' paper, the level separation (2–3) is held fixed at 7.17 kMc. The magnetic field strength and direction are varied to bring the (1–2) separation into coincidence or near coincidence with the (2–3) separation. Harmonic 1 : 1 coincidence occurs at about 22.3°. Levels 2 and 3 are saturated with a short pulse, and the decay of their populations is then observed. From the decays observed at a series of angles centered about 22°, the cross relaxation as a function of energy imbalance can be determined.

The details of the decay function calculated from the rate equations (see Ref. 4) depend rather critically on the assumptions one makes about the spin-lattice probabilities. Assuming equal s–l probablities W_L between all pairs of levels, the observed signal $S(t)$ is given by

$$S(t)\sim 1.5\exp(-t/\tau_1)+0.5\exp(-t/\tau_2), \quad (1a)$$

$$\tau_1^{-1}=1.5W_{CR}+4W_L, \quad (1b)$$

$$\tau_2^{-2}=4W_L. \quad (1c)$$

If one assumes s–l probabilities proportional to the energy separation of the levels one obtains

$$S(t)\sim 0.8\exp(-t/\tau_1)-0.3\exp(-t/\tau_2)+1.5\exp(-t/\tau_3), \quad (2a)$$

$$\tau_1^{-1}=2.17W_L, \quad (2b)$$

$$\tau_2^{-1}=6.28W_L, \quad (2c)$$

$$\tau_3^{-1}=1.5W_{CR}+2.23W_L. \quad (2d)$$

W_L is here taken as the s–l probability between levels 3 and 4.

When the relaxation process is dominated by cross relaxation, one obtains reasonable estimates of the cross relaxation rate regardless of detailed assumptions about s–l probabilities. It is clear, for example, that Eqs. (1a) and (2a) coalesce for times of the order of W_{CR}^{-1}, when W_{CR}/W_L becomes large.

Fortunately, many of the processes observed by Mims are indeed dominated by cross relaxation. One can fit the decay traces either by a double exponential [Eq. (1a)] or by a triple exponential [Eq. (2a)] in the region where cross relaxation is rapid. For cross-relaxation times shorter than a few msec, closely agreeing values result whether one uses equal, unequal, or infinite lattice times. The sensitivity of these particular experiments does not appear sufficient to allow consistent and quantitative interpretation of the residual trace after the fast component has been extracted.

TABLE I. C–R probabilities from Mims experiment.

ω(Mc)	τ(ms)	W_{CR}(sec^{-1})
−970	3.1	220
−670	2.3	290
−370	1.2	540
230	0.035	1900
530	5.3	125
830	22.	30

* This work was supported in part by the U. S. Army Signal Corps. under Contract DA-36-039-sc-39169.

[1] W. J. C. Grant, preceding papers, Phys. Rev. 134, A1554 and A1565 (1964).
[2] W. B. Mims and J. D. McGee, Phys. Rev. 119, 1233 (1960).
[3] P. S. Pershan, Phys. Rev. 117, 109 (1960).
[4] W. J. C. Grant, Phys. Chem. Solids (July, 1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We show in Table I the C-R probability versus the frequency separation of the (1–2) and (2–3) transitions, as derived from the Mims experiment. The same data are also presented in Fig. 1. Because of the rapid variation of the energy imbalance ω with the field direction, and because W_{CR} becomes of the same order of magnitude as W_L for large ω, the uncertainty in W_{CR} ranges from $\sim 10\%$ for small ω to a factor of 2 or so for large ω. We note that the last two points for negative ω fall outside the main sequence. This effect is sufficiently marked so that it should probably be considered real, and we shall comment on these points later.

The theoretical expression for W_{CR}, we recall, is essentially the convolution of two functions: χ, which is the power spectrum of the dipole operator, and Φ, which in the present case is itself the convolution of the $(1 \leftrightarrow 2)$ and $(2 \leftrightarrow 3)$ resonance lines. In the following computations, we have used the form of χ given in Eq. (53) of part I. This χ function was computed using matrix elements appropriate for the magnetic field parallel to the crystal axis, and is consequently not the exactly correct function. As pointed out in part I, however, the salient features of the χ function do not depend on the detailed angular functions assumed, and Eq. (53) of part I represents a close approach to the exact function, without requiring excessive numerical computation. We may then write

$$W_{CR}=[32\pi^2\langle m^2\rangle n^2/3\hbar v L\epsilon]\Psi(\omega; L\epsilon, a). \tag{3}$$

The quantity $\langle m^2\rangle$ is the square of the transition dipole matrix element, averaged over a sphere of unit radius; the quantity L is the change in dipole energy between two flipping spins, maximized with respect to angular position, and again for unit radius; v is the average volume per lattice site; ϵ is defined by the relation $\epsilon^{-1/3}\equiv r_0=$ the effective nearest-neighbor distance; n is the fractional mole concentration of Cr^{3+} in each level; a is the half-width of Φ. $L\epsilon$ is directly related to the width of χ, [Eq. (52), part I]. The numerical constants have been so chosen that Ψ is normalized for $L\epsilon=2g^2\beta^2$.

The quantities $\langle m^2\rangle$ and L are derived from matrix elements that can be calculated precisely. The quantity a can be estimated with reasonable accuracy. In the present case, as discussed in part II, it is the width of a function obtained by convoluting the $(1 \leftrightarrow 2)$ and $(2 \leftrightarrow 3)$ resonance line shapes. (These line shapes in turn can be calculated from first principles by methods closely related to our present theory.) The quantity r_0 is a free parameter, in the sense that its theoretical prediction lies altogether outside the present theory. The value of r_0 which our theory extracts from the cross relaxation data must, however, be compared with the radius of strong exchange as determined by Statz and co-workers.[5] Exchange-coupled pairs would not, in general, be expected to participate in the two-spin cross-relaxation process, with its sharply defined resonance.

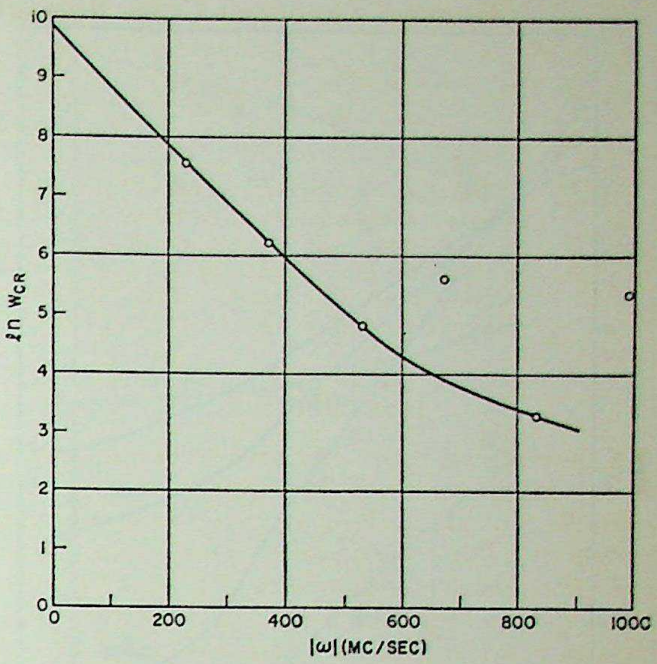


FIG. 1. Experimental data (from Mims, Ref. 2).

We first discuss the calculation of the matrix elements. These are computed from the *pair* Hamiltonian

$$\mathcal{H}_{\text{pair}}=\mathcal{H}_{\text{pair}}{}^{0}+\mathcal{H}_{\text{pair}}{}^{\text{dip}} \tag{4a}$$

$$\mathcal{H}_{\text{pair}}{}^{0}=g\beta\mathbf{H}\cdot(\mathbf{S}_1+\mathbf{S}_2)-D(S_{z1}{}^2+S_{z2}{}^2), \tag{4b}$$

$$\mathcal{H}_{\text{pair}}{}^{\text{dip}}=\frac{g^2\beta^2}{r_{12}{}^3}\left[\mathbf{S}_1\cdot\mathbf{S}_2-\frac{3(\mathbf{r}_{12}\cdot\mathbf{S}_1)(\mathbf{r}_{12}\cdot\mathbf{S}_2)}{r_{12}{}^2}\right]. \tag{4c}$$

For a particular value of $\mathbf{H}$, taking into account both its magnitude, and its direction with respect to the c axis, one diagonalizes $\mathcal{H}^0$ numerically. Then the same transformation is applied to $\mathcal{H}^{\text{dip}}$. The matrix elements relevant to a transition between the pair state (2,2) and the pair state (1,3), for example, are easily obtained from the resulting matrices. (The labels 1,2,3, refer to single-particle energy states in order of decreasing energy. They do not refer to m states, since m is not a good quantum number.) The matrix elements are complicated functions of the coordinates. For the Mims experiment they are

$$\begin{aligned}\Delta\omega\equiv&\langle(1,3)|\mathcal{H}^{\text{dip}}|(1,3)\rangle-\langle(2,2)|\mathcal{H}^{\text{dip}}|(2,2)\rangle\\&=(g^2\beta^2/r^3)[1.36(3\cos^2\theta-1)\\&\qquad+0.73\sin\theta\cos\theta+1.34\sin^2\theta]\end{aligned} \tag{5}$$

$$\begin{aligned}m\equiv&\langle(1.3)|\mathcal{H}^{\text{dip}}|(2,2)\rangle\\&=(g^2\beta^2/r^3)[0.191(3\cos^2\theta-1)\\&\qquad+1.562\sin\theta\cos\theta+1.417\sin^2\theta].\end{aligned} \tag{6}$$

These matrix elements lead to the values

$$L=2.782g^2\beta^2, \tag{7}$$

$$\langle m^2\rangle=1.425g^4\beta^4, \tag{8}$$

[5] L. Rimai, H. Statz, M. J. Weber, G. A. De Mars, and G. F. Koster, Phys. Rev. Letters 4, 125 (1960). H. Statz, L. Rimai, M. J. Weber, G. A. De Mars, and G. F. Koster, J. Appl. Phys. 32, 2185 (1961). M. J. Weber, L. Rimai, H. Statz, G. A. De Mars, Bull. Am. Phys. Soc. 6, 141 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

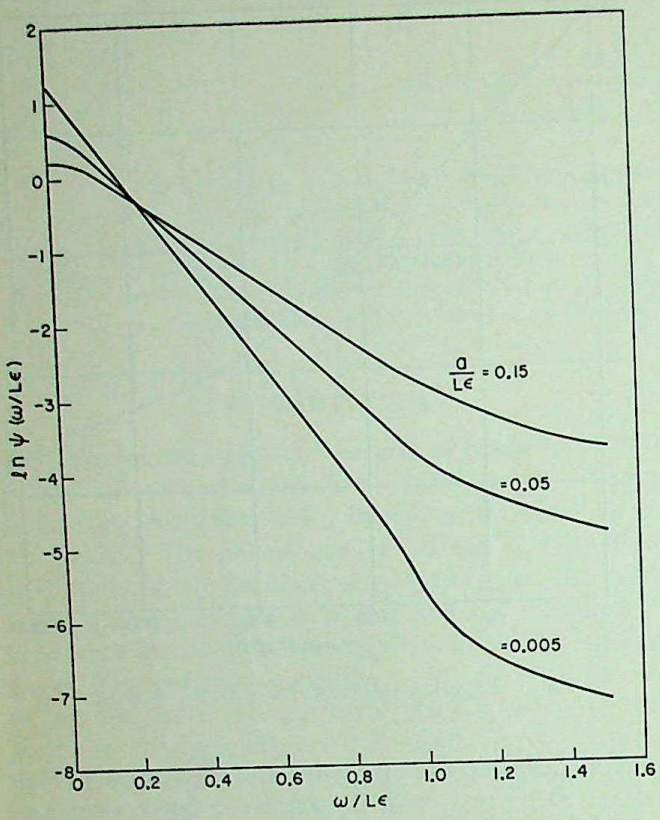


FIG. 2. ln ψ versus $\omega/L\epsilon$.

with

$$g^2\beta^2 = 5.11\times10^4 \text{ Mc(Å)}^3.$$

We note that there is no need to use projection operators. Through the device of the *pair* matrix we project the appropriate portion of the dipole operator simply by inspection of the appropriate matrix elements. Of course, the use of the pair matrix allows an exact computation of the matrix elements for any operating condition, regardless of the orientation and magnitude of **H**, that is regardless of whether or not m is a good quantum number.

We next consider the dependence of W_{CR} on a and on $L\epsilon$.

Although in principle a is determined by our knowledge of resonance line shapes, in practice the situation is complicated by the presence of inhomogeneous broadening and by the complex structure not only of the total resonance line but even of its homogeneous component. At a Cr^{3+} weight concentration of 0.05%, the half-width ascribable to homogeneous broadening is about 6 Mc, and the homogeneous line shape is a Lorentzian whose falloff gradually becomes Gaussian about 1 kMc from the center frequency. The inhomogeneous broadening is a pure Gaussian of half-width about 18 Mc. The total line is the convolution of these two. It has a half-width of about 23 Mc. At the center it is dominated by the Gaussian. At frequencies more than about 60 Mc from the center it is dominated by the Lorentzian homogeneous component. In the context of cross relaxation we are usually talking about frequency separations of the order of hundreds of Mc, where the total resonance line behaves essentially like its Lorentzian homogeneous component. Very near the exact harmonic condition, cross relaxation is so rapid that the inhomogeneous line may no longer relax as a whole, in which case we are observing relaxation of individual Lorentzian spin packets rather than of the entire system. We note that the function Φ is the convolution of the $(1\leftrightarrow2)$ and $(2\leftrightarrow3)$ resonance lines, and that in convolution, Lorentzians add width algebraically, while Gaussians add width in rms fashion. One way to overcome some of these complications is to find an approximation to Ψ [Eq. (3)] from which the dependence on $L\epsilon$ can be extracted more or less independent of detailed assumptions about a, or about the shape of the function Φ.

Fortunately, such an approximation exists. If one plots $\ln\Psi$ as a function of $\omega/L\epsilon$, one obtains curves which are very closely straight lines over several orders of magnitude in Ψ. The slope of these straight lines is dependent only weakly on the shape of Φ or on the value of a. In Fig. 2, we show $\ln\Psi$ as a function of $\omega/L\epsilon$ for various values of $a/L\epsilon$, and with Φ Lorentzian. In Fig. 3, we have plotted the quantity $(a/L\epsilon)\times d\ln W_{CR}/d(\omega/L\epsilon)$ against $a/L\epsilon$. Figure 3 is, in a sense, the meeting point between theory and experiment. If $d\ln W_{CR}/d\omega$ and a are known experimentally, then Fig. 3 determines $a/L\epsilon$, and therefore ϵ. We note that a straight line of unit slope in Fig. 3 would indicate that the slope of $\ln W_{CR}$ is independent of a. For a Gaussian Φ, the acutal curve deviates negligibly from such a straight line; for a Lorentzian Φ, it deviates only slightly for small a more seriously for large a.

Our theory then predicts that $\ln W_{CR}$ should vary almost linearly with ω over several orders of magnitude (with some rounding at the very peak of the curve). The slope of this straight line yields a close estimate of the quantity ϵ, that is, of the exchange radius r_0.

This linearity is a good approximation when $a/L\epsilon$

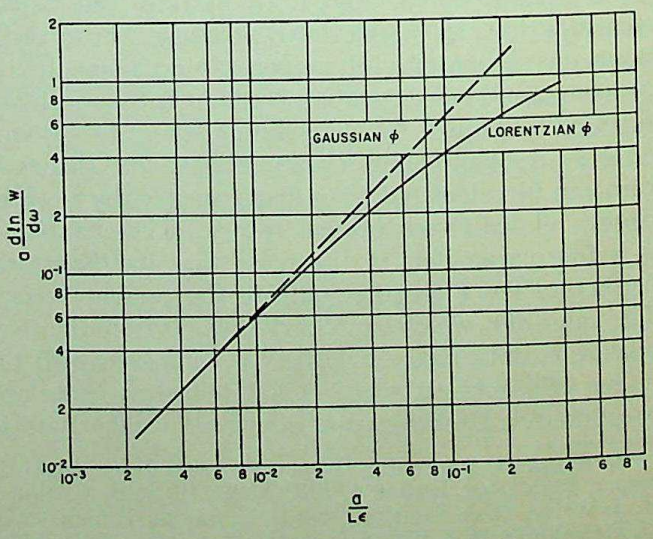


FIG. 3. $a[d(\ln\psi)/d\omega]$ versus $(a/L\epsilon)$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

is considerably smaller than unity, that is for magnetically very dilute systems. As can be seen from Fig. 1, the linear approximation to $\ln W_{CR}$ is consistent with the data of the Mims experiment.

For the Mims experiment, $d\ln W_{CR}/d\omega = 0.010$ Mc^{-1}. For a Lorentzian of half-width 12 Mc (the convolution of two Lorentzians of 6 Mc each), one obtains $r = 6.28$ Å. For a Gaussian of half-width 32 Mc (the convolution of two Gaussians of 23 Mc each) one obtains $r_0 = 6.15$ Å. In this case, therefore, it happens to make very little difference whether one uses the width of the entire resonance line or only its homogeneous component. While this is partly fortuitous, it is primarily due to insensitivity of $L\epsilon$ to the shape of Φ.

The work of Statz and co-workers[5] indicates that the exchange radius lies between about 5.5 and 6.5 Å. Thus, our value of 6.15 ± 0.15 Å is in excellent agreement with his. We feel that such corroboration is rather encouraging in view of the startling conclusion that the first twelve or thirteen neighbor shells, or about the first fifty nearest-neighbor sites, are sufficiently exchange coupled to be excluded from this cross-relaxation process.

Because of the trick of working with the derivative of $\ln W_{CR}$, we have so far concerned ourselves exclusively with the shape of $W_{CR}(\omega)$, completely separating out the question of its magnitude. The magnitude of W_{CR} depends sharply on the concentration, and the absolute concentration is difficult to determine accurately. Nominal values of concentration are accurate within about 50%. Furthermore, there is some evidence of pronounced variations in concentration, over a scale of the order of microns, which makes the value of the true effective concentration even more obscure. In view of these uncertainties, the absolute magnitude of the cross-relaxation probability calculated by our theory, as we shall now show, agrees astonishingly well with the experiments of Mims.

Since we have already accounted for the shape of $W_{CR}(\omega)$, it suffices to calculate its magnitude at one point. For convenience, we choose the peak of the curve, $W_{CR}(0)$. The required value is given by Eq. (3), using $\Psi(0; L\epsilon, a)$. For the Mims experiment, $\Psi(0; L\epsilon, a) = 2.9$, if one assumes a Lorentzian Φ. L and $\langle m^2 \rangle$ are given by Eqs. (7) and (8). The nominal weight concentration of 0.05% corresponds to a molar concentration of 0.00034 or to a concentration per level of 8.2×10^{-5} One then computes $W_{CR}(0) = 25\,900$ sec^{-1}. Extrapolation of the curve of Fig. 1, yields $W_{CR}(0) = 18\,000$ sec^{-1}.

Even leaving out of account the considerable experimental uncertainties in the determination of the points in Fig. 1, the discrepancy between these two values can be accounted for by an actual concentration 17% below nominal. One must also bear in mind that the χ function on which this analysis is based is inexact, and that truly precise numbers cannot therefore be expected. In view of these circumstances, the agreement with experiment is excellent.

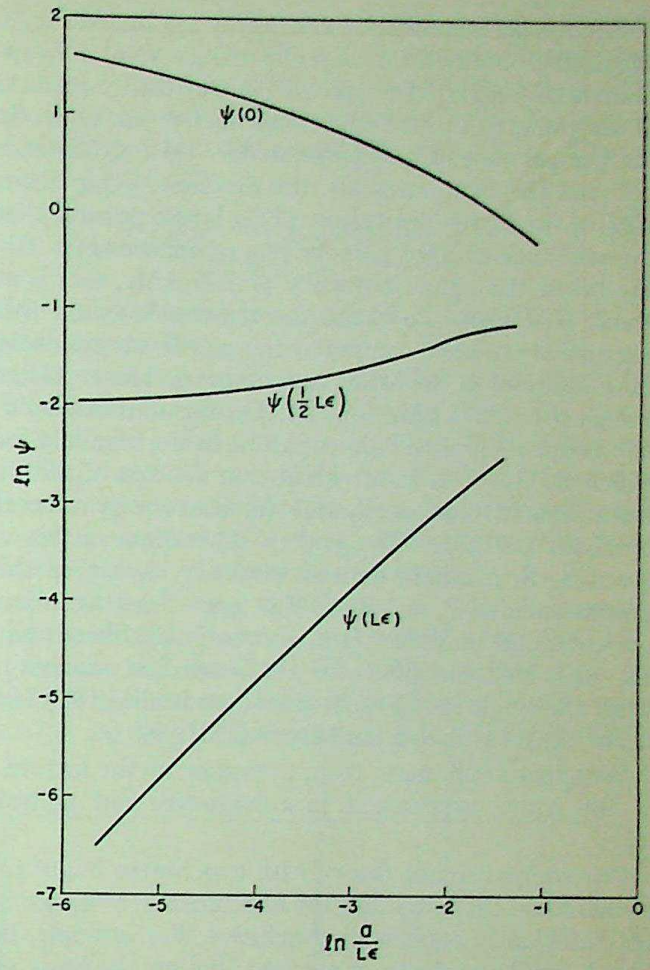


FIG. 4. $\ln\psi$ versus $\ln(a/L\epsilon)$.

We note that a linear concentration dependence of W_{CR} would give values that deviate from those observed by more than four orders of magnitude.

The interpretation of a direct measurement of the concentration dependence must be made with some care, since the concentration enters not only as a coefficient, but also determines the shape of Ψ via the parameter a. In Fig. 4 we show $\ln\psi$, evaluated at various values of ω as a function of $\ln(a/L\epsilon)$. The curves are not linear and depend drastically on the value of ω. As $a/L\epsilon$ becomes larger, the curves would all approach one another, and all would approach negative unit slope. This can be seen intuitively because ψ is normalized and approaches the normalized function Φ as $a/L\epsilon$ becomes large. The curve for $\psi(\omega = 0)$ already has a slope of -0.8 at $a/L\epsilon = 1/e$. For 0.05% ruby, $a/L\epsilon$ is about 1/20, and the slope of $\psi(0)$ is about -0.5. Since a varies almost linearly with the concentration n in this region, one would expect to measure a dependence of $n^{1.5}$ for the peak value of the cross-relaxation probability, the dependence becoming less steep at higher concentrations. Such a prediction is somewhat sanguine, since it depends not only on the precise form of χ and Φ but also on the precise values of a and $L\epsilon$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Finally, we comment on the "stray" points in Fig. 1, which are obtained for negative values of ω. Two explanations suggest themselves: (1) Our theory predicts an asymmetry in general, if $|\Delta\omega\,\max| \neq |\Delta\omega\,\min|$. In the present case [Eq. (5)] there does exist a difference between the two extremals, the difference being about 10% of the extremum values. (2) It is also possible that the exchange-coupled pairs have a resonance about 800 Mc below the signal frequency of 7.17 kMc, that is at about 6.37 kMc. To check the reasonableness of this suggestion we have computed the pair levels assuming the conditions of the Mims experiment and an exchange energy J of 28.75 kMc (or $5D$). The antiferromagnetically coupled pairs will almost all be in the manifold for which $\mathbf{S}_1\cdot\mathbf{S}_2=9/4$. It turns out that two out of the six transitions within this manifold have an energy interval of about 6.4 kMc. The energy separations within a given $\mathbf{S}_1\cdot\mathbf{S}_2$ manifold depend only very feebly on the precise value of J, as long as J is large. For this reason our computation indicates a reasonable likelihood that the pairs will contribute to the single ion relaxation when the single ion $(1\leftrightarrow 2)$ transition is about 800 Mc below the cross-relaxation harmonic frequency.

We appear to be able, then, to explain all the features of the Mims experiment in a consistent and unified way.

Our comparison of theory with experiment might be criticized on the basis that the experimental evidence is not sufficiently copious or conclusive. For example, in Fig. 1, we have drawn a straight line on the basis of three points—an extrapolation which admittedly lacks uniqueness. Nevertheless, we can assert minimally that our calculations are consistent with the experimental data available. These data include independently the cross-relaxation decay studies, the available information regarding the structure and concentration dependence of the absorption resonance, the independent determination of the exchange radius. The area of consistency extends not only to the shapes of the functions, not only to parametric dependence, but, what is most important, to actual magnitudes. We are thus actually able to tie together, in a fully quantitative way, a substantial amount of dispersed experimental evidence.

II. PERSHAN'S EXPERIMENTS ON LiF

We next consider the cross-relaxation process in LiF, which has been investigated experimentally and theoretically by Pershan,[3] in a notable study which established decisively the existence of cross relaxation. Pershan interpreted his experiments in terms of moment theory and the convolution of Gaussian resonance lines. We shall see that our own theory leads in some circumstances to similar conclusions as Pershan's; but certain experimental features which proved anomalous in terms of his formalism find their natural place in ours.

Cross relaxation in LiF contrasts sharply in several respects with cross relaxation in ruby: (1) The process of main interest is a three-spin process involving two Li's flipping opposite one F, which requires slightly more complicated rate equations. Furthermore, neglecting quadrupole interactions, there is no zero-field splitting. The 4 Li levels are equally spaced, and transitions occur simultaneously between all pairs of adjacent levels. (2) There is no exchange interaction in LiF. This frees us from some of the major concerns in ruby, namely the extent of the exchange radius and the possible influence of exchange-coupled pairs. On the other hand, as we have seen in part II, the dipole interaction itself is now dominated by nearest, or by near, neighbors. This means that the use of a χ function based on a continuous dipole density is not justified. To obtain the correct χ we must sum over the discrete sites. (3) LiF is a magnetically concentrated system. The shape of Φ is Gaussian, and there is no question of inhomogeneous broadening.

The three-spin process can occur in any of nine mutually exclusive ways, since each of the two Li atoms can be in any one of three initial states. In addition, a second-order two-spin process can occur in which a single Li spin changes its z component by two magnetons. Two initial states are possible for this process. The decay constants for each of these relaxation modes are obtained from rate equations in Ref. 4, where it is also shown that the observed decay constant is their average.

The calculation of the matrix element governing each of these decay modes has been discussed in general in part II and with special reference to LiF in Pershan's paper.[3] We recall that each matrix element is a second-order element consisting of a sum running over six terms. The intermediate states correspond to the possiblities of any one of the three spins flipping first, and the converse. The actual computations gain considerably in simplicity and lucidity by evaluating the matrix elements of products of spin operators directly in the appropriate pair manifold.

The function $\chi(\omega)$ is given by an expression analogous to Eq. (48) of part II:

$$\chi(\omega)=2\pi\sum_{pq}M^2(\mathbf{r}_{1p},\mathbf{r}_{1q})\times\delta[\Delta\omega_\chi(\mathbf{r}_{1p})+\Delta\omega_\chi(\mathbf{r}_{1q})+\Delta\omega_\chi(\mathbf{r}_{pq})-\omega]. \quad (9)$$

An Li is here taken as the reference atom 1; p labels all other Li atoms; and q labels the F atoms. The $\Delta\omega_\chi(\mathbf{r}_{ij})$ are the changes in dipole energy induced by the flipping of atoms i and j, as discussed in detail in part II. As discussed in part II, correlations are explicitly included through the term $\Delta\omega_\chi(\mathbf{r}_{pq})$. M^2 is the square of the transition matrix element we have just discussed. It consists of a series of terms typically of the form

$$(1/\Delta E)^2[m_1^2(\theta_{1p})m_2^2(\theta_{1q})/r_{1p}{}^6r_{1q}{}^6]. \quad (10)$$

The symbols m^2 and ΔE have been defined and discussed in part II, Eqs. (41)–(45), and (28). In the present case

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$\Delta E = g_{Li}\beta H$ when an Li flips first, $\Delta E = g_F\beta H$ when an F flips first.

We have computed the pq sums in Eq. (9) twice, once including the first neighbor shell only, and then the first two neighbor shells. The first shell consists of the nearest 6 F atoms and 12 Li atoms. The second shell contains the next 8 F atoms and 6 Li atoms. There are thus 72 terms in the first sum, 252 in the second. Each of the calculations has been done for the [001], [011], and [111] directions. It turns out that for the [001] direction the second shell already contributes only a small correction. For the other directions, the sum is also converging, but clearly has not yet reached its limit. Nevertheless, the main features of interest to us already appear clearly, as we shall see, from the summations over the first and second shells. For this reason we feel that the very steep escalation in computation time and labor, which would be necessary to carry the sum further, is unwarranted.

The appropriately weighted average of the χ functions is presented in Fig. 5, for each of the three-crystal directions. The units along the ω axis are gauss, with the correspondence $\omega = (2g_{Li} - g_F)\beta H$ or 1 G = 0.697 kc. The δ functions of Eq. (9) have been smoothed by averaging over 0.6-G intervals. The ω^{-2} dependence arising from the energy denominator has been suppressed. The reason is that this factor is excluded from the convolution integral for the transition probability. That is

$$W_{CR}(\omega) \sim \frac{1}{\omega^2} \int \chi'(\omega')\Phi(\omega - \omega')d\omega', \qquad (11)$$

where χ' is ω^2 times the χ defined in Eq. (9), or in Eq. (46b) or part II.

The crucial point in Fig. 5 is that χ for the [111] direction is narrowly peaked around the origin, but χ for the [001] direction is spread out to a distance of 24 G. In part II we pointed out that the existence of large components of $\chi(\omega)$ far from the origin will result in an apparent shift of the C–R shape function away from the point of harmonic coincidence.

In the final step, convoluting χ with a Gaussian, the standard deivation σ used was 13 G for the [111] direction, 14 G for the [011] direction, and 19 G for the [001] direction. Pershan obtained a best fit using 12.7 G for the [111] direction and 14.8 G for the [011]

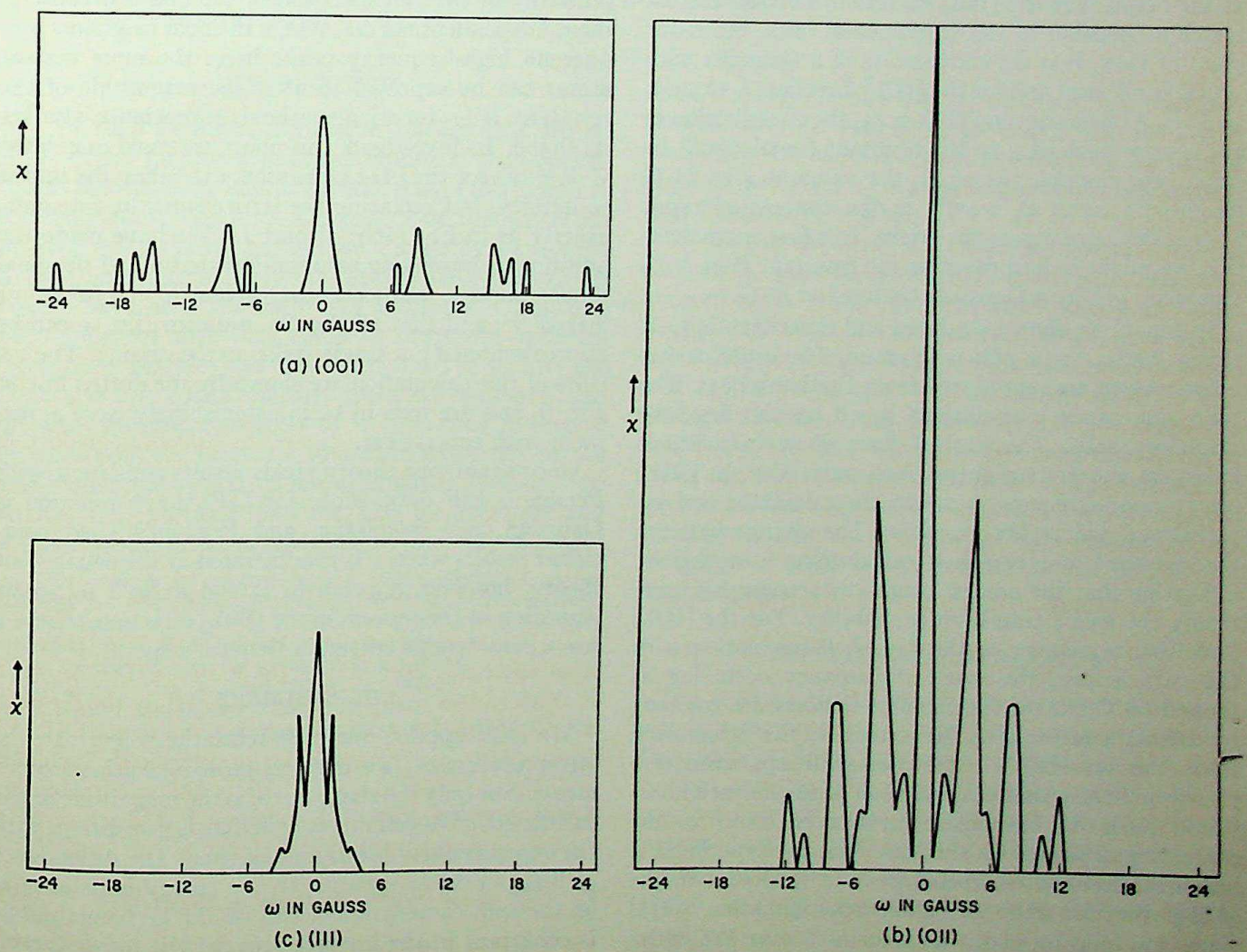


FIG. 5. $\chi(\omega)$ for LiF.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

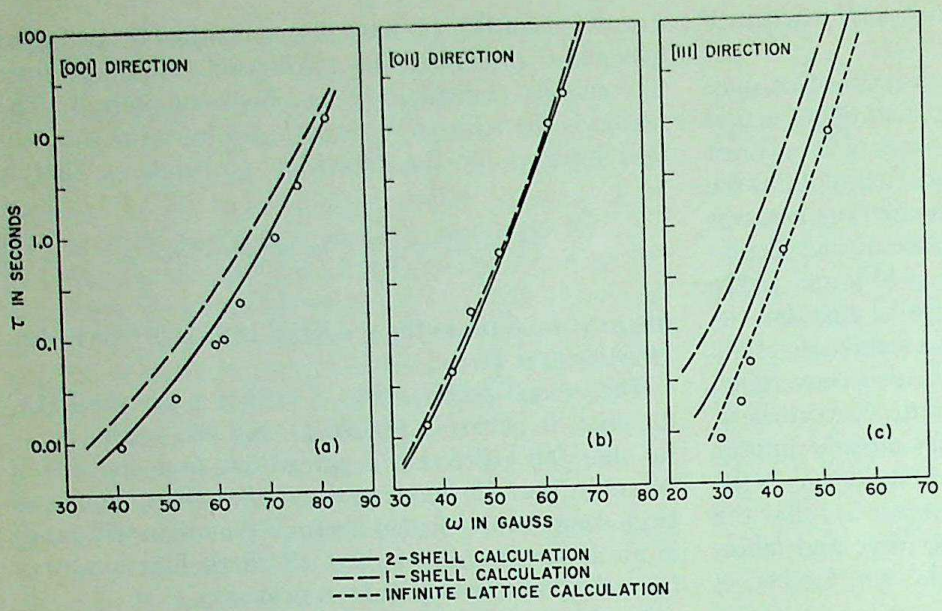


FIG. 6. C–R times for LiF.

direction. This slightly larger σ for the [011] direction compensates for his implicit approximation of χ by a δ function at the origin. For this direction, our χ does not vanish until 12 G, although it has an enormous peak at the origin. For the [001] direction, Pershan cannot match a Gaussian to the experimental data. From our point of view, it is the convolution of a Gaussian with χ that is relevant and for the [001] direction, χ extends to 24 G. At fields beyond 20 G or so, the contribution of the central peak of χ to the resultant function will be small. For example, at 65 G, the value of χ at 24 G weights 34 times as heavily in the convoluted result as does the value of χ at the origin. To a first approximation we might indeed describe the resultant W_{CR} by a Gaussian, but by a Gaussian centered at 24 G.

In Fig. 6, we show calculated and experimental C–R times. Again, the ω axis is in gauss. The experimental points shown are reproduced from Pershan's data. The solid lines show calculations based on the first two neighbor shells. The dashed lines show calculations based on the first neighbor shell only. For the [011] orientation, agreement between the calculated and experimental values are excellent. The change between the one-shell and two-shell calculations is negligible, indicating that the nearest neighbors account for practically the entire transition probability. For the [001] direction, agreement of the two-shell calculation with the data is fair, the worst discrepancy occurring at around 60 G, where the calculated times are too long by almost a factor of 2. Nevertheless, the progression from the one-shell to the two-shell approximation clearly indicates that the inclusion of further neighbors would push the calculated curve even closer to the experimental points. In this case too, the experimental numbers decisively corroborate our theory. For the [111] direction, the two-shell approximation yields times too long by about a factor of 2 over the entire observed range; the one-shell calculation yields times too long by about a factor of 10. This is to be expected. The χ function for this case consists essentially of a single sharp peak at the origin.

The contribution of the more distant neighbors is primarily in the neighborhood of the origin because of the r^{-3} dependence of $\Delta\omega$. Where the near neighbors produce no high-frequency peaks in χ, the more distant atoms can be expected to alter the magnitude of the resultant W_{CR} function, without appreciably altering its shape. To investigate this point, we have calculated C–R times for the [111] direction, extending the lattice to infinity, but replacing the lattice sums by integrals, exactly as in Eq. (46b) of part II. We have made the additional simplifying assumptions that all off-diagonal dipole matrix elements are given by $g_{\mathrm{Li}}^2\beta^2/r^3$ or $g_{\mathrm{Li}}g_{\mathrm{F}}\beta^2/r^3$, and that all the $\Delta\omega_\chi$ are zero (i.e., χ can be approximated by a single spike at the origin). The results of this calculation are shown by the dotted line in Fig. 6, and are seen to be in astonishingly close agreement with experiment.

Once again, our theory yields results consistent with Pershan's LiF data. Since for LiF, the Φ function is Gaussian, our calculation and Pershan's yield coincident results when χ is concentrated at the origin. Our theory, however, appears to extend as well to an explanation of the process in the [001] direction, where χ has a considerable spread in frequency space.

III. SUMMARY

We have applied our cross-relaxation theory to the interpretation of two classical cross-relaxation experiments. Not only the shape but also the magnitude of the calculated cross-relaxation function is consistent with the experiments of Mims and Pershan. The dependence of W_{CR}, in dilute systems, on the energy imbalance ω, on the spin concentration, and on the exchange radius is confirmed in the interpretation of the Mims experiment. The dependence of W_{CR}, in concentrated systems,

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

on interactions with near neighbors, and its associated spread and displacement in frequency is confirmed in the interpretation of Pershan's experiment.

IV. COMPUTATIONAL NOTE

This study has involved a sizeable mass of numerical detail. The generation and manipulation of large pair matrices, the tabulation of the various χ functions, the evaluation of different convolution integrals, to mention only the more obvious computational tasks, have had to be performed on a mass-production basis. The use of the Bell Telephone Laboratories IBM-7090 computer has proved indispensable. The computer programs were written by the author.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Molecular Field Model and the Magnetization of YIG*

ELMER E. ANDERSON
U. S. Naval Ordnance Laboratory, White Oak, Maryland
and
The University of Maryland, College Park, Maryland
(Received 7 January 1964)

The susceptibility and magnetization of YIG of high purity have been measured from 4.2 to 650°K by means of a precision vacuum balance. The spontaneous magnetization has a saturation value of 37.90 emu/g at 4.2°K and 27.40 emu/g at 292°K. The Curie point is at 559°K as determined by both the vanishing of the spontaneous moment and the discontinuity in the susceptibility curve. Using a program written for the IBM 7090, the molecular field coefficients were determined by fitting the experimental total magnetization curve. The sublattice magnetizations and the exchange interactions are calculated and compared with other results. On the basis of the molecular field model the intrasublattice interactions must be larger than previously supposed.

INTRODUCTION

ACCURATE sublattice magnetization data are required for an adequate description of such quantities as the magnetic anisotropy and magnetostriction in ferrimagnetics. The first calculations of the sublattice magnetizations in the garnets were made by Pauthenet in 1957 based on the molecular field model.[1] Though his results have been the only ones available it has long been apparent that they could be improved upon by using purer samples and a better method of solving the molecular field equations. An alternative approach is to measure the magnetization of a sublattice indirectly by observing the magnetic resonance frequency or the Mössbauer absorption of nuclei situated in that sublattice. This was first done for YIG by Solomon[2] and Robert[3] and has been repeated by several others,[4–9] some of whose results will be discussed in a later section. It may be said here, however, that the NMR measurements have not been extended to sufficiently high temperatures to represent any real improvement over Pauthenet's results.

In this present study the total spontaneous magnetization was obtained by subtracting the field-dependent magnetization from the measured values for high-purity YIG over the temperature range from 4.2 to 650°K. These values of the spontaneous magnetization were fed into a program written for the IBM 7090 by Gerhard Heiche of this laboratory and the molecular field equations were solved for all temperatures. The molecular field coefficients computed in this manner are used to calculate the sublattice magnetizations and the exchange interaction energies for YIG. Similar results for the three sublattice garnets will be the subject of a separate paper.

EXPERIMENTAL APPARATUS

The technique used for the measurements reported here is a modification of the Curie method.[10] Briefly, the sample is placed in a magnetic field having a large gradient and the force on the sample is measured by a sensitive balance. The essential unit is an automatic vacuum balance and recorder which can weigh accurately to 3×10^{-5} g. A quartz sample holder is suspended from one pan of the balance so that the sample

* This work is an excerpt from a thesis submitted to the University of Maryland in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The research was performed at the U. S. Naval Ordnance Laboratory.

[1] R. Pauthenet, Ann. Phys. (Paris) 3, 424 (1958).
[2] I. Solomon, Compt. Rend. 251, 2675 (1960).
[3] C. Robert, Compt. Rend. 251, 2684 (1960).
[4] G. K. Wertheim, Phys. Rev. Letters 4, 403 (1960); J. Appl. Phys. 32, 110S (1961).
[5] C. Alf and G. K. Wertheim, Bull. Am. Phys. Soc. 5, 428 (1960).
[6] E. L. Boyd, L. J. Bruner, J. I. Budnick, and R. J. Blume, Bull. Am. Phys. Soc. 6, 159 (1961).
[7] S. Ogawa and S. Morimoto, J. Phys. Soc. Japan 17, 654 (1962).
[8] L. D. Khoi and M. Buyle-Bodin, Compt. Rend. 253, 2514 (1961).
[9] E. L. Boyd, V. L. Moruzzi, and J. S. Smart, J. Appl. Phys. 34, 3049 (1963).
[10] P. Curie, Ann. Chim. Phys. (7) 5, 289 (1895).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

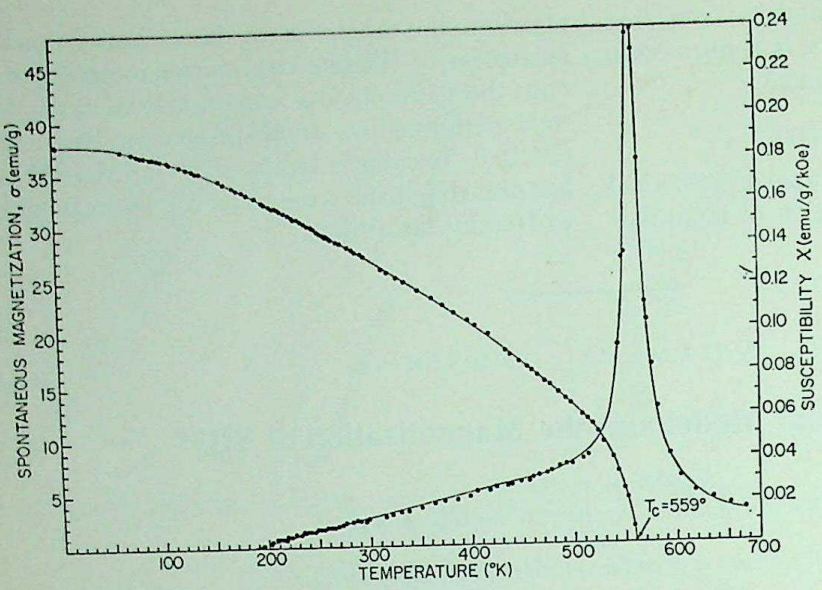


FIG. 1. Spontaneous magnetization and susceptibility of YIG.

is at the point where the magnetic field gradient is a maximum. The sample holder is protected by a vacuum-tight glass tube which may be either evacuated or subjected to a helium atmosphere at any desired pressure. For low-temperature measurements a stainless-steel double Dewar is used. High temperatures are achieved by using a noninductive wire-wound heater whose magnetoresistance coefficient is extremely low. The temperature of the sample is monitored by two thermocouples which are mounted close to the sample on the inside of the protective tube. In the interval from 50 to 250°K the maximum temperature error is estimated at ±0.5°C and above room temperature it is probably less than ±0.2°C. The magnetic field is provided by an electromagnet having parabolic pole pieces situated so as to produce a large gradient in the vertical direction. The maximum field attainable is 10 kG and the corresponding gradient is 1241 G/cm at the optimum sample position. The latter was determined by mapping the field of the magnet and choosing the point where $\partial^2 H/\partial z^2=0$. Ideally, if $\partial H/\partial z$ is the only nonzero derivative, then the force per gram of sample is given by

$$F=\sigma(\partial H/\partial z). \tag{1}$$

The gradient was determined for each value of magnet current from measurements on spectroscopically pure nickel (Johnson Matthey Code 890) using the value of 54.39 emu/g for nickel at 20°C.[11] The precision of this calibration is better than ±0.1% and includes any errors due to sample weight, sample position, and field current setting. It also includes any uncertainties due to sample shape since spherical and nonspherical samples were used to make the calibration runs.

SAMPLES

YIG single crystals were grown from the melt by J. Richard Cunningham, Jr., of this laboratory using oxides of 99.99% purity. Polycrystalline samples were made by ceramic techniques using ultrapure yttrium oxide (Lindsay Code 1118). The single crystals were ground into spheres and were x-ray oriented so that the applied field could be directed along a [111] axis. Measurements on the polycrystalline samples were made on sintered powders as well as on spheres ground from toroids of high density. X-ray diffraction patterns were obtained for all of the polycrystalline materials using a Norelco high-angle goniometer spectrometer. The scanning rate was $\frac{1}{4}$° per minute using unfiltered chromium radiation. None of the samples showed any unwanted phases or impurities and the powders appeared to be homogeneous and completely reacted.

MOLECULAR FIELD THEORY OF YIG

On the basis of the two sublattice model for ferrimagnetics [12] the molecular field equations for YIG may be expressed as follows:

$$\begin{aligned}
&\sigma=\sigma_d-\sigma_a,\\
&\sigma_a=75.70B_{5/2}(3.359\times10^{-4}/T)H_a,\\
&\sigma_d=113.55B_{5/2}(3.359\times10^{-4}/T)H_d,\\
&H_a=\lambda_{aa}\sigma_a-\lambda_{ad}\sigma_d,\\
&H_d=-\lambda_{ad}\sigma_a+\lambda_{dd}\sigma_d.
\end{aligned} \tag{2}$$

In Eqs. (2), σ is the total magnetization, σ_a is the magnetization of the octahedral sublattice and σ_d is the magnetization of the tetrahedral sublattice, all expressed in emu/g. $B(x)$ is the Brillouin function, T is the absolute temperature, and $H_{a,d}$ is the molecular field in gauss. The molecular field coefficients λ_{ij} are related to the n_{ij} of Pauthenet[1] by

$$\lambda_{ij}=92.25n_{ij}.$$

They may also be expressed in terms of the three

[11] R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951), p. 867.

[12] L. Néel, Ann. Physik 3, 137 (1948).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

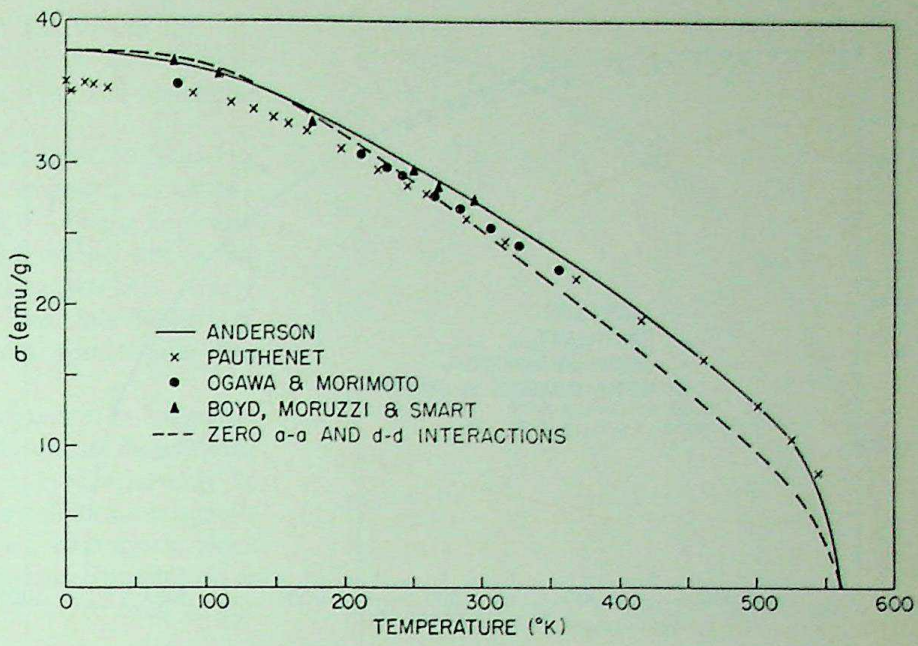


FIG. 2. Spontaneous magnetization of YIG. The solid curve is the experimental curve of Fig. 1 with data points omitted. The broken curve is the molecular field solution using the coefficients in line 2 of Table I.

exchange energies as follows:

$$\lambda_{aa}=5661J_{aa},$$
$$\lambda_{dd}=1887J_{dd}, \qquad (3)$$
$$\lambda_{ad}=2821J_{ad},$$

where the J_{ij} are in units of cm^{-1}.

Equations (2) were solved by means of a computer program designed to determine the molecular field coefficients by fitting the experimentally determined total magnetization curve.

EXPERIMENTAL RESULTS

The spontaneous magnetization and the susceptibility of YIG are shown in Fig. 1 in units of emu/g. The value of 37.90 emu/g at absolute zero corresponds to slightly over 5 Bohr magnetons and is in excellent agreement with the value of 5.01 reported by Geller *et al.*[13] This is about 5% higher than Pauthenet's value and attests to the purity and homogeneity of these samples. The Curie temperature is 559°K and is indicated by the discontinuity in the susceptibility curve as well as the vanishing of the magnetic moment. There is little question about the shape of the magnetization curve as is evidenced by the smoothness of the data. The points shown in Fig. 1 represent only about one-fourth of those measured and the omitted points fit the curve just as smoothly. The total magnetization is again shown in Fig. 2, where it is compared with the data of Pauthenet and the NMR results. The latter, unfortunately, do not extend much above room temperature and thus shed no light on the region where the magnetization changes most rapidly. The broken curve in Fig. 2 is the molecular field solution obtained by assuming that the *a–a* and *d–d* interactions are zero. Adjusting the *a–d* interaction coefficient to give the experimental Curie temperature of 559°K, one obtains the results shown in line 3 of Table I. However, the experimental curve (solid line) of Fig. 2 is best fit by the two sets of coefficients in lines 1 and 2 of Table I. The coefficients of line 1 were obtained by assuming them to be independent of temperature while those of line 2 were given Aléonard's temperature dependence, namely,

$$\lambda(T)=\lambda(o)\cdot(1-1.3\times10^{-4}T).$$

The magnetization curve calculated from each of these sets of coefficients is extremely close to the experimental values. The relative error is generally less than a few tenths of a percent except near the Curie point. Two factors probably contribute to the latter: (1) The spontaneous magnetization in this region is quite small and the measured magnetization consists largely of the field-dependent contribution. Thus, a small error in the susceptibility will produce a large relative error in spontaneous moment. (2) The molecular field model neglects short-range order which is known to play an ever increasing role as the long-range order disappears.

TABLE I. Molecular field coefficients for YIG (J_{ij} in cm^{-1}).

	λ_{ad}	λ_{aa}	λ_{dd}	J_{ad}	J_{aa}	J_{dd}	T_c °K
Anderson	71 505	47 820	22 394	25.36	8.45	11.86	559
Anderson	71 505	44 200	20 700	25.36	7.8	11.0	559
Anderson	38 371	0	0	13.61	0	0	559
Aléonard[a]	71 100	34 500	19 500	25.2	6.1	10.3	612
Pauthenet[b]	68 500	32 500	19 700	24.3	5.7	10.4	600
Wojtowicz[c]	68 500	0	0	24.3	0	0	>900

[a] Reference 14. [b] Reference 1. [c] Reference 15.

[13] S. Geller, H. J. Williams, R. C. Sherwood, and G. P. Espinosa, Phys. Chem. Solids 23, 1525 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

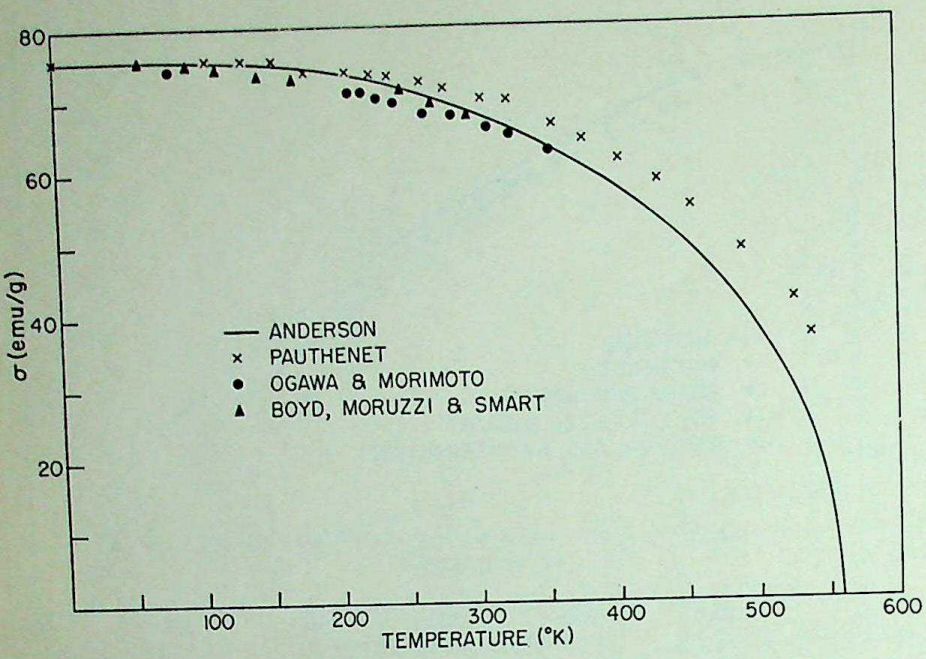


FIG. 3. Magnetization of the octahedral sublattice of YIG. The solid curve was calculated using the coefficients in line 1 of Table I.

DISCUSSION

It is evident from Fig. 2 that the shape of the total magnetization curve is not very sensitive to the values of the a–a and d–d interactions provided the a–d interaction is constrained to satisfy the Curie temperature. Thus, scattered data can be approximately fitted by a wide range of values for the coefficients. The curve in Fig. 1, however, is sufficiently smooth to give the unambiguous sets of molecular field coefficients of lines 1 and 2, Table I. These coefficients and the calculated exchange energies are compared with earlier values in the table. The tabulated values for the Curie temperature are all calculated values which were obtained from the solution of Eqs. (2) using the corresponding sets of coefficients. Note that the coefficients of lines 1–3 are the only ones that give a Curie temperature which agrees with the experimental value of 559°K. Pauthenet obtained a Curie temperature of 560°K experimentally but his temperature-dependent coefficients predict a value 40° higher.

Applying the results of spin-wave theory[16–18] to YIG, the dispersion constant D and the Landau-Lifshitz exchange constant A are related to the J's as follows:

$$\omega\hbar = Dk^2 = (a^3A/4S)k^2 = \tfrac{5}{16}(5J_{ad}-8J_{aa}-3J_{dd})a^2k^2. \quad (4)$$

Using $a=12.376\times10^{-8}$ cm and $S=\frac{5}{2}$,

$$D=4.787\times10^{-15}(5J_{ad}-8J_{aa}-3J_{dd}),$$
$$A=2.525\times10^{7}(5J_{ad}-8J_{a1}-3J_{dd}).$$

A summary of the values of D and A that have been obtained by magnetization, specific heat[19–22] and microwave measurements[23,24] is given in Table II. The values calculated from lines 1 and 2, Table I are smaller than the earlier magnetic results and are much less than those obtained from specific heat measurements. The discrepancy with the other magnetic results arises

TABLE II. Summary of exchange coefficients for YIG.

	$5J_{ad}-8J_{aa}-3J_{dd}$	D(erg-cm²)	A(erg/cm)
Magnetization			
Line 1, Table I	4.69×10^{-15} erg	22.5×10^{-30}	11.8×10^{-8}
Line 2, Table I	6.24	29.9	15.8
Line 3, Table I	13.5	64.6	34.1
Pauthenet	8.9	43.	22.
Aléonard	9.2	44.	23.
Wojtowicz	24.1	115.	61.
Specific heat			
Edmonds and Petersen		51.	27.
Meyer and Harris		63.	33.
Kunzler, Walker, and Galt		83.	44.
Shinozaki		83.	44.
Microwave			
Turner		99.	53.
LeCraw and Walker		96.	52.

[14] R. Aléonard, Phys. Chem. Solids 15, 167 (1960).
[15] P. J. Wojtowicz, J. Appl. Phys. 33, 1257 (1962).
[16] C. Herring and C. Kittel, Phys. Rev. 81, 869 (1951).
[17] R. L. Douglass, Phys. Rev. 120, 1612 (1960).
[18] H. Meyer and A. B. Harris, J. Appl. Phys. Suppl. 31, 49S (1960).
[19] D. T. Edmonds and R. G. Petersen, Phys. Rev. Letters 2, 499 (1959).
[20] J. E. Kunzler, L. R. Walker, and J. K. Galt, Phys. Rev. 119, 1609 (1960).
[21] S. S. Shinozaki, Phys. Rev. 122, 388 (1961).
[22] A. B. Harris and H. Meyer, Phys. Rev. 127, 101 (1962).
[23] E. H. Turner, Phys. Rev. Letters 5, 100 (1960).
[24] R. C. LeCraw and L. R. Walker, J. Appl. Phys. Suppl. 32, 167S (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

principally from the differences in the values for J_{aa}. The disagreement with the specific heat data is probably due to differences in the spin wave and molecular field models.

Figures 3 and 4 show the octahedral and tetrahedral sublattice magnetizations for all temperatures as calculated using the coefficients of line 1 (solid line) and line 3 (dashed line) of Table I. The dashed line is not shown for the octahedral sublattice since it nearly coincides with the solid curve. The calculations of Pauthenet and some of the NMR results are also shown.

In Mössbauer and NMR experiments in magnetic materials it is assumed that the temperature dependence of the effective field at the nucleus is due entirely to the temperature dependence of the sublattice magnetization. This is probably not a bad assumption since it has been shown[25] that the hyperfine interaction is the dominant contributor to the effective field at the nucleus.

The hyperfine term of the Hamiltonian may be written in the form

$$H = -g_I \beta_I \mathbf{I} \cdot \mathbf{H}_{\text{eff}}, \tag{5}$$

where

$$\mathbf{H}_{\text{eff}} = (8\pi/3)\rho(0) g\beta \mathbf{S} \tag{6}$$

is the effective field seen by the nuclear spin. Since $\rho(0)$ is nonzero only for s electrons, it is not obvious why Fe^{3+}, which has no unpaired s electrons, should have an effective field at the nucleus. Sternheimer[26] suggested that a net unpaired spin density at the nucleus is produced by a polarization of the core electrons due to the spin of the d electrons. If this core polarization is essentially constant up to the Curie temperature, then H_{eff} and the hyperfine energy will show the same temperature dependence as the sublattice magnetization. That this assumption is valid, at least up to 430°K in YIG, is borne out by the NMR data in Figs. 3 and 4.

SUMMARY OF RESULTS

(1) The saturation moment of YIG is 37.90 emu/g (5.01 Bohr magnetons) at absolute zero and 27.40 emu/g (3.62 Bohr magnetons) at 292°K. The Curie temperature is 559°K. The magnetic susceptibility vanishes below 190°K.

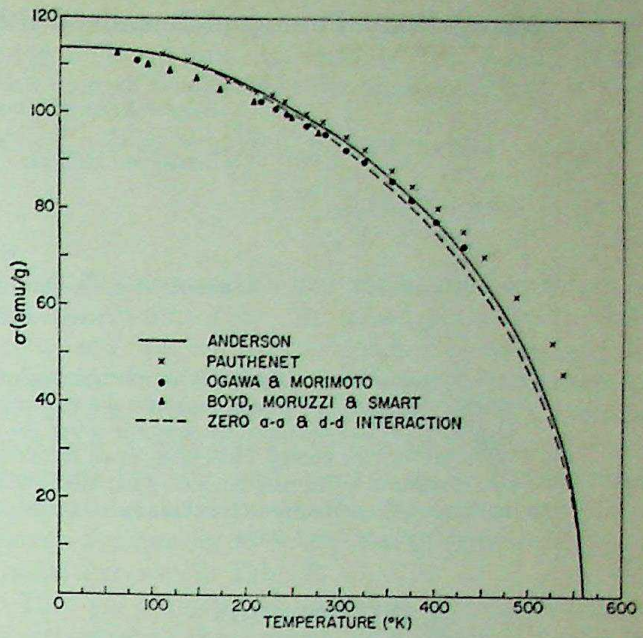


FIG. 4. Magnetization of the tetrahedral sublattice of YIG. The solid curve was calculated using the coefficients in line 1 of Table I, the broken curve using line 3.

(2) The experimental magnetization curve can be reproduced by means of the molecular field model. However, a unique set of interaction coefficients cannot be obtained by this method if there is much scatter in the data.

(3) The sublattice magnetizations are less sensitive than the total magnetization to the exact values of the interaction coefficients.

(4) The sublattice magnetizations reported here are in good agreement with available NMR results.

(5) On the basis of this model the a–a interaction must be larger than previously supposed.

ACKNOWLEDGMENTS

I am especially grateful to Professor Ralph D. Myers for his guidance during the course of this work. I wish to also thank J. Richard Cunningham, Jr., for growing the crystals, and Gerhard Heiche for writing the computer program for the solution of the molecular field coefficients.

Note added in proof. A. B. Harris [Phys. Rev. **132**, 2398 (1963)] gives the following values for the molecular field coefficients of YIG: $\lambda_{ad} = 89\,700$, $\lambda_{aa} = 35\,950$ and $\lambda_{dd} = 11\,980$. However, this set of coefficients results in a Curie temperature that is greater than 900°K.

[25] R. E. Watson and A. J. Freeman, Phys. Rev. **120**, 1125 (1960); J. Appl. Phys. Suppl. **32**, 118S (1961).
[26] R. Sternheimer, Phys. Rev. **86**, 316 (1952).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Many-Body Theory of Localized d States in Metals. II. Pair of Impurities*

P. GOTTLIEB
Hughes Research Laboratories, Malibu, California
and
University of California, San Diego, La Jolla, California

AND

H. SUHL
University of California, San Diego, La Jolla, California

(Received 20 January 1964)

A previous rotationally invariant perturbation treatment of localized d states is extended to the case of neighboring sites interacting via direct d-d exchange and indirect s-d scattering. It is found that the parallel alignment of two localized moments is never the ground state of the pair. For certain values of the unperturbed d-state energy (depending upon the exchange interactions) only one of the sites will be occupied by a d electron so the pair can have a net moment. If the conditions are such that both sites are occupied by d electrons, these electrons must be aligned antiparallel so there is only a small net moment due to scattering to and from the band states.

I. INTRODUCTION

IN a previous paper,[1] a rotationally invariant perturbation calculation of the susceptibility of a localized d state in a metal was made, using a model of this state proposed by Anderson.[2] In this model the "d state" is assumed to have no orbital degeneracy, and to have a wave function distinct from the Wannier functions of electrons on the sites of the host lattice. When this state is singly occupied, its energy is ϵ_d, when doubly occupied by electrons with antiparallel spins, is $2\epsilon_d+U$, where U is positive. The Hamiltonian of the model consists of the band electron energies, the aforementioned "d state" energies, and a scattering term capable of converting electrons on the local "d" site into band electrons and conversely. No other terms (such as Coulomb interaction between conduction electrons and electrons on the "d" site) are included. Physical reasons for restriction of the energies to the ones mentioned above, are given by Anderson.[2] Within the framework of this model, he showed that, provided the scattering to the band was not too strong relative to the correlation energy U, and provided ϵ_d was below the Fermi surface and $2\epsilon_d+U$ above, a localized magnetic moment would result, which could be calculated by the Hartree-Fock method. Since this method violates rotational invariance, it is hard to see how one can use it to compute a Curie constant as distinct from a saturation magnetization.

In SFI it was shown how this calculation may be done by perturbation theory. The method divided the possible states of the system into magnetic and nonmagnetic ones, and calculated two partition functions $\mathfrak{Z}_M$, $\mathfrak{Z}_S$ in these two manifolds. The Curie constant then has the form

$$C_{\text{eff}}=C_{1/2}[\mathfrak{Z}_M/(\mathfrak{Z}_M+\mathfrak{Z}_S)].$$

The actual calculation used the resolvent technique, deriving resolvents $\mathfrak{R}_M$ and $\mathfrak{R}_S$ for the two manifolds with the help of projection operators. These resolvents were then evaluated by perturbation theory, keeping only the lowest order self-energy corrections and, in fact, only their imaginary parts. The crucial point is the following: If these imaginary parts (due, of course, to the band scattering) were neglected, the partition function $\mathfrak{Z}_S$ of the nonmagnetic manifold would be smaller than $\mathfrak{Z}_M$ by a factor $e^{\beta\epsilon_d}$, for the "empty" component of the nonmagnetic manifold, and by a factor $e^{-\beta(2\epsilon_d+U)}$ for the "filled" component of that manifold. (Energies are measured from the Fermi level.) Hence, at very low temperatures, the magnetic moment would be essentially that of the isolated "d" state without scattering to the band. The imaginary parts of the self-energies appropriate to $\mathfrak{Z}_S$, which are in effect the damping rates of the states of S due to the presence of M, cause $\mathfrak{Z}_S$ to assume the *same* exponential dependence on the temperature as that of $\mathfrak{Z}_M$. For this reason there is always a reduction of the moment. Qualitatively, the results agreed with those of Anderson, except no sharp cutoff was found in the moment, which tended to zero smoothly with increasing scattering. This might be a basic feature of perturbation theory, unless prescriptions are found for solving for the self-energies consistently in the way in which this is done in other many body problems. The complicating feature here, which makes the standard recipes hard to apply, is that the "d"-state correlation energy is treated rigorously as part of the unperturbed propagator.

In the case of two d states, we distinguish between two situations: (1) The d sites are nearest neighbors and electrons on them interact directly, and (2) the d sites are far apart and interact only via band electrons.

In case (1), we must include in the Hamiltonian an

* Supported in part by the U. S. Air Force Office of Scientific Research.

[1] H. Suhl and D. R. Fredkin, Phys. Rev. 131, 1063 (1963) (hereafter called SFI).

[2] P. W. Anderson, Phys. Rev. 124, 41 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

essentially single-particle scattering between electrons on the two "d" sites. In addition, we will show that the direct Coulomb interaction between electrons on these sites may become important even if smaller than the former, provided its exchange part has ferromagnetic sign. To these interactions we must, of course, add the same Hamiltonians as in SFI, one for each site. In case (2), all direct d interaction is neglected and only the sum of the two Hamiltonians of SFI remains.

Problem (1) has also been treated in Hartree-Fock approximation by Alexander and Anderson,[3] who conclude that the pair aligns antiparallel with zero net moment unless one of the energies ϵ_d or $2\epsilon_d+U$ is close to the Fermi level. In that event, they can obtain parallel coupling, and therefore a net moment. As we shall see, perturbation theory parallels these conclusions somewhat.

II. RESOLVENT METHOD

We will treat the "d" pair Hamiltonian rigorously, and, as in SFI, only the band scattering as a perturbation. (A simultaneous treatment of both these mechanisms will be briefly considered at the end of this paper.) Then we may divide the unperturbed states of the system into three manifolds, broadly describable as singlet, doublet, and triplet, and the effective Curie constant, by a trivial generalization of the formula in SFI as[4]

$$C_{\text{eff}}=\frac{C_1\mathfrak{Z}_T+C_{1/2}\mathfrak{Z}_D}{\mathfrak{Z}_T+\mathfrak{Z}_D+\mathfrak{Z}_S}, \tag{1}$$

where C_1 and $C_{1/2}$ are the "bare" Curie constants appropriate to a free spin one and a free spin $\frac{1}{2}$, respectively, and $\mathfrak{Z}_T$, $\mathfrak{Z}_D$, $\mathfrak{Z}_S$ are the partition functions of the three manifolds. These manifolds, split into submanifolds, which may be classified according to their origin before the single particle scattering between the two sites is turned on. The result is as follows:

(S) One electron on each site, the pair forming a singlet. Two antiparallel electrons on one site, the other site empty. Both sites doubly occupied by antiparallel electrons. Both sites empty.

(D) One electron on either site, the other empty. Two antiparallel electrons on one site, the other singly occupied.

(T) One electron on each site, the pair forming a triplet.

Including the possible S_z states in D and T, this gives a total of 16 *unperturbed basis states*. When the single-particle scattering between the two sites is turned on, some of these form admixtures. The d part of the Hamiltonian (exclusive of the band scattering) is

$$\mathfrak{H}_d=\sum_\sigma \epsilon_d(n_{1\sigma}+n_{2\sigma})+U(n_{1\uparrow}n_{1\downarrow}+n_{2\uparrow}n_{2\downarrow}) +V_{12}\sum_\sigma(d_{1\sigma}{}^*d_{2\sigma}+d_{2\sigma}{}^*d_{1\sigma}), \tag{2}$$

where d^*, d are creation and annihilation operators on the sites indicated by the subscripts, σ is the spin index and the n's are occupation numbers on the indicated sites. Coulomb interation is neglected for the present. Under the V_{12} term in Eq. (2), the submanifolds enumerated above (linearly combined where required) combine and regroup as in the following table. In the ket symbols, the first and second places denote occupancies of sites 1 and 2, respectively. The results are given in Table I.

The low-temperature behavior of the partition functions of *all* these manifolds is dictated by that of the partition function of the energetically lowest manifold. (Just as in SFI, the others are "damped" by lowest manifold and so acquire its temperature dependence.) For this reason we need only determine which energy is lowest as the parameters ϵ_d, U and V_{12} are varied.

We assume $\epsilon_d<0$, $2\epsilon_d+U>0$ so that the isolated sites favor single occupancy let $V_{12}>0$. If $V_{12}>\frac{2}{3}U$, the $E(S_{m1})$ is always lowest. If $V_{12}<\frac{1}{3}U$, then $E(S_{m1})$ or $E(D_1{}^+)$ is lowest according to whether

$$-\epsilon_d \gtrless V_{12}\{1-8(V_{12}/U)[1+(1+(16V_{12}{}^2/U^2))]^{-1}\}.$$

If $V_{12}<0$, the same result holds, but with V_{12} replaced by $|V_{12}|$ and $D_1{}^+$ by $D_1{}^-$. Presumably for *reasonably* small V_{12}, the situation with singlet lowest is more common. Only when ϵ_d approaches the Fermi level to within $\sim V_{12}$, can the doublet energy become lowest.

It follows that in the common situation, any moment exhibited by the pair is due to band scattering. For small values of the scattering energy V, such moment as there is comes principally from $\mathfrak{Z}_D$, since the doublet manifold can reach the singlet manifold by lowest order damping involving V^2. The triplet manifold connects to the singlet only in order V^4. Qualitatively, this agrees with the Hartree-Fock procedure of Alexander and Anderson, who find no moment at all unless $|V_{12}|\sim|\epsilon_d|$.

The triplet energy is *always* above $E(S_{m1})$. It is also below $E(D_1{}^+)$, unless

$$V_{12}>-\epsilon_d>V_{12}\{1-8(V_{12}/U) \times[1+(1+(16V_{12}{}^2/U^2))^{1/2}]^{-1}\},$$

in which case $E(D_1{}^+)$ lies between the singlet S_{m1} and the triplet. In that case the triplet partition function acquires the exponential temperature dependence of that of the doublet, but at temperatures such that kT is much less than a quantity of order $V_{12}{}^2/U$, the damping process still favors the singlet.

[3] S. Alexander and P. W. Anderson, Phys. Rev. 133, A1594 (1964).

[4] It should be emphasized again that formulas like (1) and (7a) of SFI give *only* the thermal expectation value of $n_{d\uparrow}-n_{d\downarrow}$ in the case of (7a) SFI, and of the total moment on the pair of d sites alone in the case of (1) of the present paper. Any polarization effects of the conduction electrons over and above their background Pauli moment are not determined by these formulas.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Energies of various states.[a]

States	Energies
Singlets	
both sites empty, $S_e=\|0;0\rangle$	0
both sites filled, $S_f=\|\uparrow\downarrow;\uparrow\downarrow\rangle$	$4\epsilon_d+2U$
"polar" singlet; $S_p=(2)^{-1/2}(\|\uparrow\downarrow;0\rangle-\|0;\uparrow\downarrow\rangle)$	$2\epsilon_d+U$
"mixed" singlets	
$S_{m1}=(2)^{-1/2}[a(\|\uparrow;\downarrow\rangle-\|\downarrow;\uparrow\rangle)+b(\|\uparrow\downarrow,0\rangle+\|0,\uparrow\downarrow\rangle)]$	$2\epsilon_d-U\sinh^2(\theta/2)$
$S_{m2}=(2)^{-1/2}[-b(\|\uparrow;\downarrow\rangle-\|\downarrow;\uparrow\rangle)+a(\|\uparrow\downarrow,0\rangle+\|0,\uparrow\downarrow\rangle)]$	$2\epsilon_d+U\cosh^2(\theta/2)$
Doublets	
one-electron doublets; $D_1^{\pm}=(2)^{-1/2}(\|\uparrow,0\rangle\pm\|0,\uparrow\rangle)$	$\epsilon_d\mp\frac{1}{4}U\sinh\theta$
three-electron doublets: $D_3^{\pm}=(2)^{-1/2}(\|\uparrow\downarrow,\uparrow\rangle\pm\|\uparrow,\uparrow\downarrow\rangle)$	$3\epsilon_d+U\mp\frac{1}{4}U\sinh\theta$
and two similar pairs with the spins reversed.	
Triplets	
$\|\uparrow;\uparrow\rangle$	$2\epsilon_d$
$T=(2)^{-1/2}(\|\uparrow;\downarrow\rangle+\|\downarrow;\uparrow\rangle)$	$2\epsilon_d$
$\|\downarrow;\downarrow\rangle$	$2\epsilon_d$

[a] In this table $\sinh\theta=4V_{12}/U$, $a=-\coth(\theta/2)/[1+\coth^2(\theta/2)]^{1/2}$, and $b=1/[1+\coth^2(\theta/2)]^{1/2}$. For brevity, the energies of these states S_f, S_p, etc. will be denoted by $E(S_f)$, $E(S_p)$, etc.

Finally, when $E(D_1^+)$ is also below $E(S_{m1})$ (see inequality), the susceptibility for small band scattering is dominated by spin $\frac{1}{2}$ with V^2 corrections from spin 1, and spin 0. This is the situation in which Alexander and Anderson find triplet coupling favored. In the present theory no such coupling ever occurs.

However, some caution is indicated by the fact that in the present model the singlet-triplet separation is only of order V_{12}^2/U. If a Coulomb interaction between the electrons on the d sites is added, and its exchange energy W, say, turns out to have ferromagnetic sign, the above conclusions become invalid as soon as W exceeds V_{12}^2/U. The triplet would then tend to be lowest, and the susceptibility would be dominated by spin one.

III. EVALUATION OF THE PARTITION FUNCTIONS

We use the resolvent technique (as in SFI) for evaluation of the partition function. The complete Hamiltonian is taken to be

$$\mathcal{H}=\sum_{k,\sigma}\epsilon_k c_{k,\sigma}^{\dagger}c_{k,\sigma}+V\sum_{k,\sigma}\sum_{i=1}^{2}(e^{+i\mathbf{k}\mathbf{r}}c_{k,\sigma}^{*}d_{i,\sigma}+e^{-i\mathbf{k}\mathbf{r}}d_{i,\sigma}^{*}c_{k,\sigma})+\mathcal{H}_d$$
$$\equiv\mathcal{H}_b+\mathcal{H}_d+\mathcal{V}=\mathcal{H}_0+\mathcal{V}, \tag{3}$$

where $\mathcal{H}_d$ is the Hamiltonian of the "d" states alone, $\pm\mathbf{r}$ are the position vectors of the sites, and $\mathcal{V}$ is the band scattering operator. The resolvent operator $\mathfrak{R}=(\mathcal{H}-z)^{-1}$, could be projected onto the unperturbed singlet, doublet, and triplet manifolds as in SFI. The result will involve "nested" propagators and have an extremely complicated appearance, each propagator involving the manifold one more remote than its predecessor. We have calculated one of these nested expressions in the approximation in which, at most, two conduction electrons or holes are excited in the intermediate states.

The resolvent for the D_1 state (which exhibits the behavior of resolvents of other states as well), with the denominator evaluated to first order in Δ ($\Delta=\pi\rho V^2$, $\rho=$density of states), may be written as

$$\mathfrak{R}_{D_{1(z)}}=[E(D_1)-z+\sum_k \mathcal{T}_k+i\delta]^{-1}, \tag{4}$$

where δ is an infinitesimal,

$$\mathcal{T}_k=\frac{|\langle D_1|\mathcal{V}|T\rangle|^2 n_k}{E(T)-z-\epsilon_k}+\sum_{S'}\frac{|\langle D_1|\mathcal{V}|S'\rangle|^2 n_k}{E(S')-z-\epsilon_k}+\frac{|\langle D_1|\mathcal{V}|S_e\rangle|^2(1-n_k)}{-z+\epsilon_k},$$

and $\sum_{S'}$ means the sum over the three singlet states with two d electrons, S_p, S_{m1}, S_{m2}. The matrix elements may be evaluated with the aid of the wave functions derived above. For example, $N|\langle D_1^+|\mathcal{V}|S_{m1}\rangle|^2=\gamma_+V^2\times\cos^2\mathbf{k}\cdot\mathbf{r}$, $N|\langle D_1^+|\mathcal{V}|T\rangle|^2=3V^2\sin^2\mathbf{k}\cdot\mathbf{r}$, where $\gamma_{\pm}=1\pm[4V_{12}/(U^2+16V_{12}^2)^{1/2}]$. The summation over k is somewhat complicated by the presence of $\sin^2\mathbf{k}\cdot\mathbf{r}$ (etc.). Since ϵ_k is isotropic, the angular integration will give

$$\int d\Omega\begin{Bmatrix}\sin^2\mathbf{k}\cdot\mathbf{r}\\ \cos^2\mathbf{k}\cdot\mathbf{r}\end{Bmatrix}=2\pi\begin{Bmatrix}1-\dfrac{\sin 2kr}{2kr}\\ 1+\dfrac{\sin 2kr}{2kr}\end{Bmatrix}=2\pi\begin{Bmatrix}\mu_-\\ \mu_+\end{Bmatrix}.$$

If we assume

$$|\epsilon_k(dk/dE)_F|\ll\max(k_F,1/r),$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

then

$$\frac{\sin 2kr}{2kr}=\frac{\cos 2k_Fr\sin 2[\epsilon_k(dk/dE)_Fr]+\cos 2[\epsilon_k(dk/dE)_Fr]\sin 2k_Fr}{2[k_F+\epsilon_k(dk/dE)_F]r}\approx\frac{\sin 2k_Fr}{2k_Fr}.$$

Then $\mu_\pm$ becomes

$$\alpha_\pm=1\pm(\sin 2k_Fr/2k_Fr).$$

Now we will neglect the real part of $\sum_k\mathcal{T}_k$, and calculate

$$\begin{aligned}(2/\Delta)\mathrm{Im}\sum_k\mathcal{T}_k=&3n[E(T)-z]\alpha_-\\&+n[E(S_p)-z]\alpha_-+n[E(S_{m1})-z]\alpha_+\gamma_-\\&+n[E(S_{m2})-z]\alpha_+\gamma_++2n[-z]\alpha_+,\end{aligned}\quad(5)$$

where $n[z]$ is the Fermi distribution. Following the procedure of SFI, the partition function may be written as

$$\begin{aligned}\mathfrak{Z}_{D1}{}^+&=\int_{-\infty}^{+\infty}\mathrm{Im}\mathfrak{N}(z)e^{-\beta z}dz\\&=e^{-\beta E(D_1{}^+)}\int_{-\infty}^{+\infty}\frac{[\delta+\Delta I(\zeta)]e^{-\beta\zeta}d\zeta}{\zeta^2+[\Delta I(\zeta)+\delta]^2},\end{aligned}\quad(6)$$

where

$$I(\zeta)=\frac{1}{\Delta}\mathrm{Im}\sum_k\mathcal{T}_k$$

and we have substituted $\zeta+E(D_1{}^+)=z$.

$I(\zeta)$ will be infinitesimally small unless $\zeta>E(S_{m1})-E(D_1{}^+)$, and the integrand of Eq. (6) will be infinitesimally small unless $\zeta<0$. The problem then requires consideration of two separate cases depending on the sign of $E(S_{m1})-E(D_1{}^+)$.

A. Magnetic: $E_0=E(S_{m1})-E(D_1{}^+)>0$

In this case the integrand behaves like a delta function near $\zeta=0$, in the limit $\delta\to0$. The $I(\zeta)$ term in the numerator can only contribute for $\zeta>E_0$, and, since $e^{-\beta E_0}\ll1$, this contribution is negligible. (For large negative values of ζ the integrand could become large because of the exponential factor, even compensating for the small factor δ. However, either the band bottom or the finite spatial extent of the impurity potential V will introduce a cutoff before this happens.) The partition function for the magnetic case is now simply

$$\mathfrak{Z}_{D_1}{}^+=e^{-\beta E(D_1{}^+)}.\quad(7)$$

B. Nonmagnetic: $E_0=E(S_{m1})-E(D_1{}^+)<0$

$I(\zeta)$ is now finite at $\zeta=0$ so there is no contribution from the δ in the numerator in the limit $\delta\to0$. If we make the transformation

$$e^{-\beta\zeta}n(\alpha-\zeta)=e^{-\beta\alpha}n(\zeta-\alpha),$$

then we immediately see that the largest term in $e^{-\beta\zeta}I(\zeta)$ is

$$\tfrac{1}{2}e^{-\beta E_0}\alpha_+n(\zeta-E_0)\gamma_-.$$

Hence, the partition function may be written

$$\mathfrak{Z}_{D_1}{}^+=\gamma_-e^{-\beta E(S_{m1})}\frac{\Delta}{2}\int_{-\infty}^{E_0}\frac{\alpha_+d\zeta}{\zeta^2}.$$

The $I(\zeta)$ does not contribute to the denominator because we are integrating over that portion of the ζ axis where $I(\zeta)=0$. The straightforward evaluation of the integral then gives

$$\mathfrak{Z}_{D_1}{}^+=-(\Delta/2E_0)\gamma_-\alpha_+e^{-\beta E(S_{m1})}.\quad(8)$$

Calculation of Curie constant. We could proceed to evaluate all the other partition functions in the same manner. However, examination of Eqs. (7) and (8) and the analogous results in SFI will enable us to write the following set of simplifying rules for the calculation of the partition functions:

(1) All partition functions have the form $n_iA_ie^{-\beta E_l}$, where A_i is independent of temperature, $E_l=\min\times[E(S_{m1}),E(D_1{}^+)]$, and n_i is an integer indicating the multiplicity of the state.

(2) The partition function for the state of lowest energy has $A=1$.

(3) For any state coupled to the lowest state to first order in $\mathcal{U}$ (states coupled in higher order are neglected in the present paper),

$$A_i=(\Delta\gamma_\pm\alpha_\pm)/(E_i-E_l).$$

The α and γ factors are determined from the square of the appropriate coupling matrix element.

Applying these rules, we can calculate the effective Curie constant from Eq. (1) for the two cases.

Nonmagnetic: $E_0=E(S_{m1})-E(D_1{}^+)<0$

$$C_{\mathrm{eff}}=C_{1/2}F/(1+F),\quad(9)$$

where

$$F=\frac{\Delta}{2}\left[\frac{\alpha_+\gamma_-}{E(D_1{}^+)-E(S_{m1})}+\frac{\gamma_+\alpha_-}{E(D_1{}^-)-E(S_{m1})}+\frac{\gamma_+\alpha_+}{E(D_3{}^+)-E(S_{m1})}+\frac{\gamma_-\alpha_-}{E(D_3{}^-)-E(S_{m1})}\right].$$

The labelings of the energies in the denominators of F indicate the four levels to which S_{m1} is coupled in first order.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Magnetic: $E_0=E(S_{m1})-E(D_1^+)>0$

$$C_{\rm eff}=\left\{C_{1/2}+\frac{3C_1\Delta\alpha_+}{2[E(T)-E(D_1^+)]}\right\}\times\left\{1+\frac{\Delta}{2}\left[\frac{\gamma_+\alpha_+}{E(S_{m2})-E(D_1^+)}+\frac{\gamma_-\alpha_+}{E(S_{m1})-E(D_1^+)}+\frac{\alpha_+}{E(S_p)-E(D_1^+)}+\frac{3\alpha_+}{E(T)-E(D_1^+)}\right]\right\}^{-1}. \quad (10)$$

Each doublet state is actually counted twice, but since all the terms represent doublets, the factors of two cancel out. The accuracy of these expressions requires $\Delta\ll|E(S_{m1})-E(D_1^+)|$. In the region of transition from magnetic to nonmagnetic, the two levels cross and this inequality cannot hold. We must therefore do a more detailed calculation involving only the S_{m1} and D_1^+ levels.

IV. ZERO TEMPERATURE LEVEL SPLITTING

When the S_{m1} and D_1^+ levels are close together, we are permitted to rediagonalize them with respect to $\mathcal{U}$ and ignore the other levels. In this section we will therefore assume $|E(D_1^+)-E(S_{m1})|\ll|E(T)-E(D_1^+)|$, $|E(S_{m2})-E(D_1^+)|$, etc. For $E(S_{m1})<E(D_1^+)$ it is appropriate to consider a two-electron manifold. We take the wave function for this restricted manifold to be of the form

$$\psi=\sum_{k,\sigma}a_{k,\sigma}|D_1^+\rangle c_{k,-\sigma}+b|S_{m1}\rangle, \quad (11)$$

with the coefficients normalized so that

$$\sum_{k,\sigma}|a_{k,\sigma}|^2(1-n_k)+|b|^2=1.$$

The renormalized eigenvalues are found by minimizing the energy [calculated with Eq. (11) as wave function], minus the normalization condition multiplied by λ, with respect to a_k and b. This procedure gives the equations

$$\begin{aligned}&[E(D_1^+)+\epsilon_k-\lambda]a_k+V(\gamma_-\alpha_+/2)^{1/2}b=0,\\ &V(\gamma_-\alpha_+/2)^{1/2}\sum_k a_k(1-n_k)+[E(S_{m1})-\lambda]b=0,\end{aligned} \quad (12)$$

determining the new energies as solutions of

$$-\lambda+E(S_{m1})=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{1-n_k}{E(D_1^+)+\epsilon_k-\lambda}. \quad (13)$$

It should be noted that we have actually renormalized only the states S_{m1} and D_1^+ accompanied by an electron. If $E(S_{m1})>E(D_1^+)$, the appropriate manifold contains only one-electron states. The wave function is then of the form

$$\psi=a|D_1^+\rangle+\sum b_k|S_{m1}\rangle c_k. \quad (14)$$

This leads to the characteristic equation

$$-\lambda'+E(D_1^+)=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{n_k}{E(S_{m1})=\epsilon_k-\lambda'}, \quad (15)$$

which renormalizes the states D_1^+ and S_{m1} accompanied by a hole.

It can be seen at once that for $|E(D_1^+)-E(S_{m1})|\gg\Delta$, we have $\lambda=E(S_{m1})$, $\lambda'=E(D_1^+)$. Using this value of λ, and the normalization condition we find

$$a_k=\frac{V(\gamma_-\alpha_+/2)^{1/2}}{(1+F')^{1/2}[E(S_{m1})-E(D_1^+)-\epsilon_k]},$$

where

$$F'=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{1-n_k}{[E(S_{m1})-E(D_1^+)-\epsilon_k]^2}.$$

If the summation is performed, it is readily seen that F' is the same as the first term of the F used in Eq. (9). With this value of a_k, we can obtain the result

$$C_{\rm eff}=C_{1/2}\sum_k a_k^2(1-n_k), \quad (16)$$

which is the same as Eq. (9), except that the terms involving levels other than S_{m1} and D_1^+ are missing. Hence, the choice of the wave function given by Eq. (11) corresponds to choosing $E(S_{m1})<E(D_1^+)$, as would be expected.

By the same reasoning, we can use $\lambda'=E(D_1^+)$ to find the coefficient a for the wave function given by Eq. (14). In this case

$$C_{\rm eff}=C_{1/2}a^2, \quad (17)$$

which is easily seen to correspond to Eq. (10), with the terms involving states other than S_{m1} and D_1^+ removed. In order to determine the energy eigenvalues λ' when the states are close together, we rewrite Eq. (15) as

$$+y'=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{n_k}{y'-\epsilon_k+E_0}, \quad (18)$$

where

$$y'=E(D_1^+)-\lambda'.$$

For positive E_0, Eq. (17) will always have one positive root, which gives the shift in energy of the D_1^+ state when it is lowest. If we approximate the density of states by a parabolic band starting at $k=0$, then the sum can be converted to an integral from $k=0$ to $k=k_F$ which can be easily evaluated. If $|y+E_0|\ll E_F$, then the resulting equation simplifies to

$$y'=3V^2\left[\frac{1}{2E_F}\ln\frac{4E_F}{y'+E_0}-\frac{1}{E_F}\right]. \quad (18a)$$

For a convenient numerical example, the parameters $3V^2/E_F^2=10^{-4}$, $E_0=0$ give the solution

$$y'=3.6\times10^{-4}E_F.$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The solution of Eq. (13) is quite similar, provided certain approximations are made. The k summation is over all electron states above the Fermi surface, so a simple parabolic density would cause the integral to diverge. This divergence is always stopped by the actual high-momentum cutoff of the interaction potential V (which has here been assumed independent of k). For purposes of simplicity we will assume the density of conduction electron states to be symmetric with respect to the Fermi surface, although this is actually inconsistent with the assumption of a parabolic band. If we change variable to $y=E(S_{m1})-\lambda$, then Eq. (13) becomes

$$y=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{1-n_k}{y+\epsilon_k-E_0}=\tfrac{1}{2}V^2\gamma_-\alpha_+\sum_k\frac{n_k}{y-\epsilon_k-E_0}.$$

For $E_0<0$ (singlet lowest) this always has a positive root, and the result of the summation is the same as Eq. (18a) with the substitutions $y'\Rightarrow y$, $E_0\Rightarrow -E_0$. Hence, in this approximation, the unperturbed crossing point remains unshifted. Using the value of y' calculated from Eqs. (18a), (16) gives

$$C_{\rm eff}=\frac{C_{1/2}}{1+[\Delta\gamma_-\alpha_+/2(y'+E_0)]},\quad E_0>0 \qquad (19a)$$

and

$$C_{\rm eff}=\frac{C_{1/2}\Delta\gamma_-\alpha_+/2(y-E_0)}{1+[\Delta\gamma_-\alpha_+/2(y-E_0)]},\quad E_0<0. \qquad (19b)$$

The only restriction on the size of Δ is that Δ be very much smaller than the "distance" between $E(S_{m1})$ or $E(D_1{}^+)$ and any of the neglected states. There is, however, a limitation on the size of the correction term since $\ln 4E_F/y>1$, Eq. (18a) shows that y, $y'>V^2/E_F\approx\Delta$ (for anything resembling a parabolic band), so the Δ terms never become larger than one. This implies that the effective moment is always larger for $E_0>0$ than for $E_0<0$, in agreement with the previous distinction between magnetic and nonmagnetic cases. In general in this order there will be a sharp discontinuity (but no longer an infinity) in the moment at the boundary between the two regimes.

Equation (18a) will permit positive solutions for y' for slightly negative E_0. These solutions are slightly larger than for small positive values of E_0. However, this does not mean that the corrected energies do not cross. The increase in y' is only logarithmic in E_0 while the unperturbed separation of energies is linear in E_0, and therefore the latter is dominant.

V. LONG-RANGE EFFECT

If the two impurities are separated by many lattice spacings, the d-d exchange interaction V_{12} is negligible. In this case, the s-d scattering will dominate, and the resolvent can be calculated with respect to a simpler set of wave functions than those used above. For the magnetic case ($\epsilon_d<0$; $\epsilon_d+U>0$) the unperturbed ground state consists of single occupancy of the d level of each site. The two electrons can form a spin singlet or triplet, and these two levels will be degenerate in energy (to zero order in V). In order to determine whether the pair is magnetic or nonmagnetic, we must examine the manner in which the s-d interaction removes the degeneracy.

In the unperturbed representation, the triplet and completely empty and filled states are the same as were used previously. The other states are:

State	Energy
$S_1=\lvert\uparrow\downarrow;0\rangle$	$2\epsilon_d+U$
$S_2=\lvert 0;\uparrow\downarrow\rangle$	$2\epsilon_d+U$
$S_m=(2)^{-1/2}(\lvert\uparrow;\downarrow\rangle-\lvert\downarrow;\uparrow\rangle)$	$2\epsilon_d$
$D_1{}^{1,2}=\lvert\uparrow;0\rangle,\lvert 0;\uparrow\rangle$	ϵ_d
$D_3{}^{1,2}=\lvert\uparrow;\uparrow\downarrow\rangle,\lvert\uparrow\downarrow;\uparrow\rangle$	$3\epsilon_d+U$

S_m and T are the degenerate unperturbed ground states. As far as s-d scattering is concerned, the only difference between the S_m and T states is that the T state (because of the exclusion principle) will not allow a scattering from the d to the conduction band and back to the d of the opposite site.

The lowest order terms to show this effect will involve coupling the S_m or T to the S_1 or S_2 via D_1 or D_3 intermediate states. For example, the interaction term for the sequence $S_m\rightarrow D_1{}^1\rightarrow S_1\rightarrow D_1{}^1\rightarrow S_m$ and $T\rightarrow D_1{}^1\rightarrow S_1\rightarrow D_1{}^1\rightarrow T$ is

$$\frac{\tfrac{1}{2}V^4\sqrt{2}\ \mathrm{Tr}(c_{\mathbf{k}''\uparrow}c_{\mathbf{k}''\uparrow}{}^*\pm c_{\mathbf{k}''\downarrow}c_{\mathbf{k}''\downarrow}{}^*)(c_{\mathbf{k}'\uparrow}c_{\mathbf{k}\uparrow}{}^*\pm c_{\mathbf{k}'\downarrow}c_{\mathbf{k}\downarrow}{}^*)\cos(\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{k}'')\cdot\mathbf{r}}{(+\epsilon_d-z+\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}'}+\epsilon_{\mathbf{k}''})(2\epsilon_d+U-z+\epsilon_{\mathbf{k}}-\epsilon_{\mathbf{k}'})(\epsilon_d-z+\epsilon_{\mathbf{k}})}, \qquad (20)$$

where the upper (+) signs refer to the S_m sequence and the lower (−) signs refer to the T sequence. Taking the trace of the c_k operator product gives

$$[n_{\mathbf{k}'\uparrow}(1-n_{\mathbf{k}\uparrow})+n_{\mathbf{k}'\downarrow}(1-n_{\mathbf{k}\downarrow})]\delta_{\mathbf{k}'\mathbf{k}''}\delta_{\mathbf{k}\mathbf{k}''}+[(1-n_{\mathbf{k}''\uparrow})\pm(1-n_{\mathbf{k}''\downarrow})][(1-n_{\mathbf{k}\uparrow})\pm(1-n_{\mathbf{k}\downarrow})]\times\delta_{\mathbf{k}\mathbf{k}'}\delta_{\mathbf{k}''\mathbf{k}'''}. \qquad (21)$$

The first term of this expression is the same for both S_m and T; hence, it does not help remove the degeneracy. It is also independent of position because the δ function conditions make the argument of the cos in Eq. (20) equal to zero. The second term will give zero in the T state unless the conduction electrons are polarized. We will therefore consider only that part of Eq. (20) which is due to the second term of Eq. (21). If we assume the density of conduction electrons to be isotropic and sum over k', k''', average the directions of k and k'', and replace k'' by k' for simplicity, then we

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

have

$$\frac{V^4}{8r^2}\sqrt{2}\frac{f^2}{2\epsilon_d+U-z},$$

where

$$f=\sum_k \frac{(1-n_k)\sin 2kr}{k(\epsilon_d-z+\epsilon_k)}.$$

(It should be noted that this sum is really no longer discrete because the angle has been averaged out.) There are three other sequences of this class, and they are obtained by substituting D_3 and D_1 in one or both places. The result of adding all four terms together is

$$\tau=\frac{V^4\sqrt{2}}{8r^2}\frac{(f+g)^2}{2\epsilon_d+U-z},$$

where

$$g=\sum_k \frac{n_k \sin 2kr}{k(3\epsilon_d+U-z-\epsilon_k)}.$$

The main contribution to the partition function integral will come from z close to $2\epsilon_d$. Both f and g are real in this region. If we make the approximations

$$k_F r\gg 1, \quad -(r\epsilon_d m/hk_F)\gg 1,$$

and neglect terms which are exponentially small in these quantities, then an asymptotic approximation gives

$$f=\frac{\rho_s \cos 2k_F r}{rm(\epsilon_d-z)}, \quad g=\frac{\rho_s \cos 2k_F r}{rm(3\epsilon_d+U-z)}.$$

Now τ only occurs for S_{m1}, and for $z\approx 2\epsilon_d$ we find $\tau>0$. Hence, the S_{m1} is lowered with respect to the T state.

VI. SIMULTANEOUS CONSIDERATION OF *d-d* EXCHANGE AND BAND SCATTERING

The above calculations seem to indicate that the singlet (S_{m1}) is always lower than the triplet (T). In order to show that this is not due to our perturbation treatment for the band scattering, we have performed the diagonalization procedure of Sec. II, with the additional inclusion of all those states coupled by V. (This will be exact in the two-electron manifold.) The trial function consists of the four possible states with two d electrons with opposite spins ($d_{1\uparrow}{}^* d_{2\downarrow}{}^*$, $d_{1\downarrow}{}^* d_{2\uparrow}{}^*$, $d_{1\uparrow}{}^* d_{1\downarrow}{}^*$, $d_{2\uparrow}{}^* d_{2\downarrow}{}^*$), the four possible states with one electron and one conduction electron ($d_{1\uparrow}{}^* c_{k\downarrow}$, $d_{2\uparrow}{}^* c_{k\downarrow}{}^*$, $d_{1\downarrow}{}^* c_{k\uparrow}{}^*$, $d_{2\downarrow}{}^* c_{k\uparrow}{}^*$, with the conduction electron averaged over k states as in Sec. IV), and finally the four possible states with three d electrons and one conduction hole ($d_{1\uparrow}{}^* d_{2\uparrow}{}^* d_{2\downarrow}{}^* c_{k\downarrow}$, $d_{2\uparrow}{}^* d_{1\uparrow}{}^* d_{1\downarrow}{}^* c_{k\downarrow}$, $d_{1\downarrow}{}^* d_{2\uparrow}{}^* d_{2\downarrow}{}^* c_{k\uparrow}$, $d_{2\downarrow}{}^* d_{1\uparrow}{}^* d_{1\downarrow}{}^* c_{k\uparrow}$, with the conduction holes averaged over k). Using the method of Sec. IV we can diagonalize the energy by means of the equations obtained from the energy by differentiating with respect to each of the twelve coefficients of the trial states (remembering that eight of these coefficients depend upon k). The solutions are rather cumbersome, but if we assume $k_F r\gg 1$, then the four states with two d electrons combine in a manner similar to that found in Sec. II. The wave functions and energy eigenvalues have the same form as S_p, S_{m1}, S_{m2}, T ($S_z=0$) with the modifications

$$V_{12}\Rightarrow V_{12}'=V_{12}\left[1+V^2\sum_k \frac{1-n_k}{(\epsilon_d+\epsilon_k-\lambda)^2-W^2} + V^2\sum_k \frac{n_k}{(3\epsilon_d+U-\epsilon_k-\lambda)-W^2}\right]$$

and

$$\epsilon_d\Rightarrow \epsilon_d'=\epsilon_d-V^2\sum_k \frac{(\epsilon_d+\epsilon_k-\lambda)(1-n_k)}{(\epsilon_k+\epsilon_k-\lambda)^2-W^2} - V^2\sum_k \frac{(3\epsilon_d+U-\lambda)n_k}{(3\epsilon_d+U-\epsilon_k-\lambda)^2-W^2}.$$

As in Sec. IV, λ is simply the corrected energy of the state under consideration. Since these corrections have the same form for both S_{m1} and T, it is apparent that they have no effect on the previous conclusion that $E(S_{m1})<E(T)$.

VII. CONCLUSION

The calculations of this paper indicate that in perturbation theory neither the *d-d* exchange, the *s-d* scattering, nor any combination of the two will produce a parallel alignment of a pair of localized moments. The analysis does, however, show that the pair can show a sizable magnetic moment provided that the *d-d* exchange is such that $E(D_1{}^{\pm})<E(S_{m1})$. The conditions required for this relation are similar to those required by Alexander and Anderson[3] for parallel alignment. This suggests that the two methods may be equivalent, even though the representation of Alexander and Anderson[3] does not characterize the parallel alignment as either doublet or triplet.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Low-Temperature Magnetoresistance in Magnesium and Aluminum Containing Small Concentrations of Manganese or Iron*

F. T. HEDGCOCK AND Y. MUTO†
The Franklin Institute Laboratories for Research and Development, Philadelphia, Pennsylvania
(Received 20 January 1964)

Magnetoresistance measurements in magnetic fields up to 21 kOe have been made on Mg-Mn, Mg-Cd, Mg-Al, Al-Mn, and Al-Fe alloys in the temperature region of liquid helium. Magnesium alloys containing more than 0.1-at.% Mn which exhibit a resistance maximum and minimum in zero field, show a negative magnetoresistance, whereas the more dilute samples (0.001–0.1 at. % Mn) show a positive magnetoresistance, the magnitude of which decreases with decreasing temperature. The magnesium alloys containing non-transition element impurities, as well as the aluminum alloys containing transition metal impurities, are found to obey Kohler's rule. From an analysis of these data it is found that the magnetoresistivity of a dilute alloy of magnesium containing manganese, can be considered as the sum of a normal positive magnetoresistivity (obeying Kohler's rule) and an anomalous term which is negative in sign, does not obey Kohler's rule and is presumably due to a magnetic scattering of the conduction electrons. Using values of the *s-d* exchange integral and the Coulomb scattering integral derived from an analysis of the zero-field resistivity permits an explanation of the magnetoresistivity based on Kasuya's theory, at temperatures near the Néel point.

1. INTRODUCTION

NUMEROUS studies made on dilute monovalent and divalent[1,2] metals containing paramagnetic impurities have indicated a negative magnetoresistance at low temperature when the alloy exhibits both a low temperature resistance minimum and maximum. Recent studies on the[3] Cu-Fe and[4] Zn-Mn known to exhibit only a minimum down to the lowest temperatures available indicate the presence of a negative magnetoresistance in the more concentrated samples. A magnesium alloy containing small concentrations of manganese has been shown[5] to exhibit deviations from the Kohler rule, but no conclusive evidence for a negative magnetoresistance was found even at the lowest temperature (1.3°K) and the highest magnetic field available (25 kOe). Magnetoresistance studies on dilute aluminum alloys containing either manganese or iron have been carried out at liquid-helium temperatures at magnetic fields of[6] 1.5 kOe and no anomalies were found. In the present paper we wish to report some magnetoresistance studies on Mg alloys known to exhibit both a resistance minimum and a maximum[7] as well as measurements on normal magnesium alloys exhibiting no resistance anomalies. Also included in the present work are studies on aluminum-manganese and aluminum-iron alloys where the magnetoresistivity is studied to fields of 21 kOe.

2. SPECIMEN PREPARATION AND EXPERIMENTAL METHOD

The dilute Mg-Mn, Al-Mn and Al-Fe alloys used in the present investigation are the same as those used earlier in the studies of the electrical and magnetic properties of these systems.[7–9] The high-concentration Mg-Mn alloys were prepared by quenching from near the melting point after the samples had received a high-temperature homogenizing anneal. The manganese concentration of each Mg alloy was estimated from the

TABLE I. Alloy number, solute concentration, resistance ratio and resistivity at 273°K for all samples studied. Sample number refers to detailed spectroscopic analysis given in Refs. 8 and 12.

Alloy or metal	Sample number	Concentration of solute (at.%)	$\rho_{4.2°K}/\rho_{273°K}$	$\rho_{273°K}$ $\mu\Omega$ cm
Pure Mg	728	...	0.00351	4.12
MgMn	87812	0.001	0.00679	4.25
	87813	0.01	0.0111	4.28
	90911	0.08	0.1231	4.54
	91081 (a)	0.35	0.3188	5.47
	91081 (b)	0.74	0.5140	6.83
MgCd	396	0.55	0.0955	4.55
	397	1.01	0.1369	4.82
MgAl	400	0.29	0.1242	4.71
	401	0.53	0.2133	5.26
	402	0.80	0.2893	5.77
Pure Al	GKP	...	0.00187	...
AlMn	GKO	0.045	0.1209	...
AlFe	GKK	0.02	0.0045	...

* This work was supported by Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio, under Contract AF 33(657)-8744 and the U. S. Office of Naval Research under Contract NONR-3644(00).

† On leave of absence from The Research Institute for Iron, Steel, and other Metals, Tohoku University, Sendai, Japan.

[1] A. N. Gerritsen, Physica 25, 489 (1959) (a tabulation of experimental results up to 1959).

[2] G. J. van den Berg and J. de Nobel, J. Phys. Radium 23, 665 (1962) (a recent review).

[3] Y. Muto, K. Noto and F. T. Hedgcock, Can J. Phys. 42, 15 (1964).

[4] Y. Muto, Sci. Rept. Res. Inst. Tohoku Univ. (A) 13, 1 (1961).

[5] R. T. Webber, Phys. Rev. 105, 1437 (1957).

[6] J. G. Thomas and E. Mendoza, Phil. Mag. 43, 900 (1952).

[7] G. Gaudet, F. T. Hedgcock, G. Lamarche, and E. W. Wallingford, Can. J. Phys. 38, 1134 (1960).

[8] F. T. Hedgcock, W. B. Muir, and E. W. Wallingford, Can. J. Phys. 38, 376 (1960).

[9] E. W. Collings and F. T. Hedgcock, Phys. Rev. 126, 1654 (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

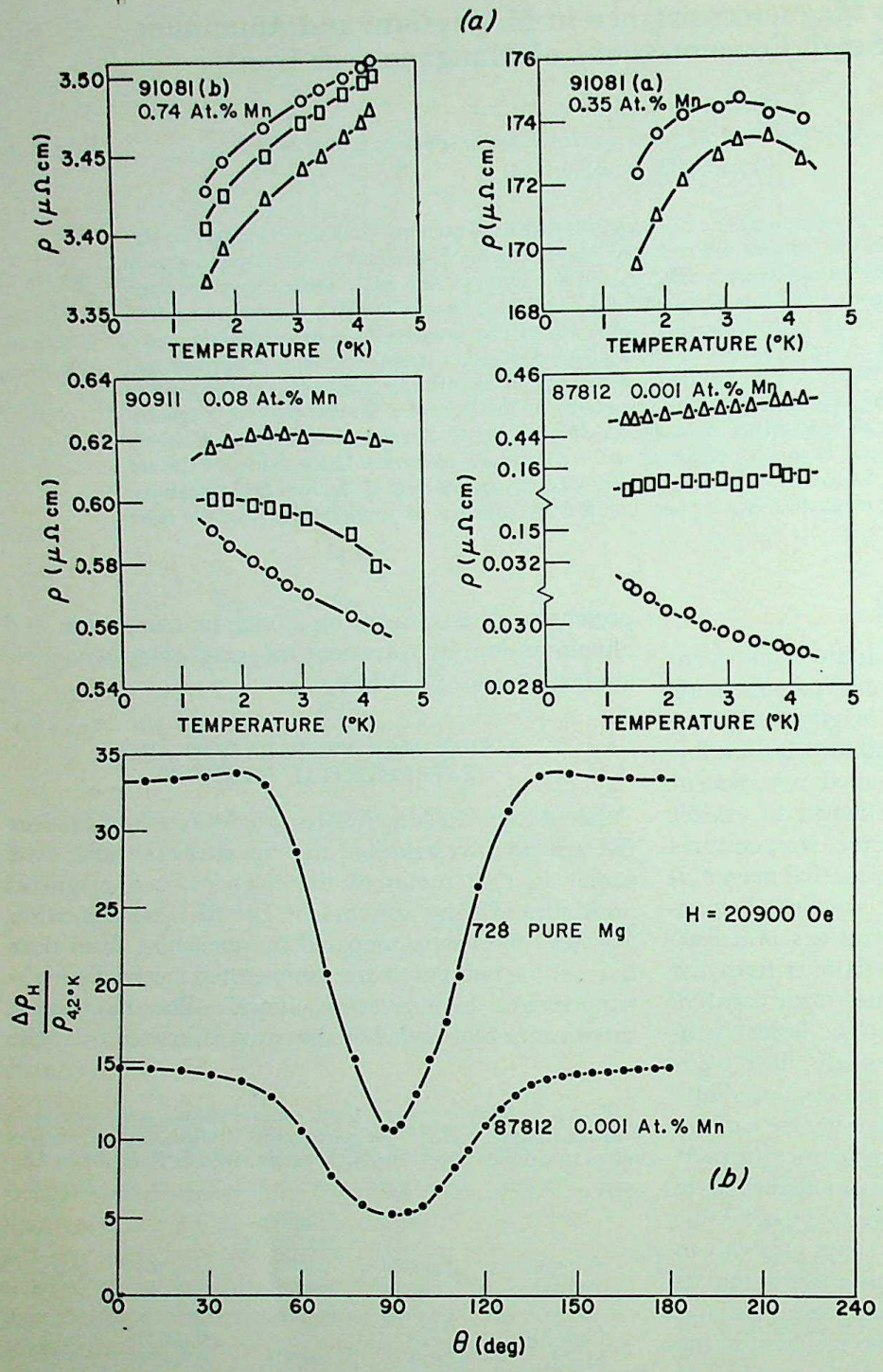


FIG. 1. (a) Resistivity of some Mg-Mn alloys as a function of temperature in a constant magnetic field. ○: 0 kOe; □: 9.93 kOe; △: 20.9 kOe. (b) Anisotropy in the magnetoresistivity at 4.2°K for pure Mg and the most dilute Mg-Mn samples. The angle axis was selected to obtain agreement in the minimum value of $\Delta\rho_{H\max}/\rho_{4.2°K}$ for both samples. $\Delta\rho_{H\max}$ means $\Delta\rho$ at the maximum field studied (20.9 kOe).

smoothed graph of resistance ratio versus nominal composition.[10] The concentration and resistivity of the Mg-Al and Mg-Cd alloys were derived from the earlier data reported by Salkovitz[11] *et al.*[12] The concentration of manganese or iron in the aluminum alloys is derived from the spectroscopic analysis. Table I contains values of the resistance ratio and resistivity for the alloys studied. The resistance samples were in the form of rolled strips (approximately $10\times0.3\times0.03$ cm); the contacts in the form of knife edges were pressed into place on a bakelite former containing the sample. A current-potential method was used for the resistance measurements with a galvanometer amplifier to read the voltages. The sensitivity of the amplifier was better than 3×10^{-9} V/mm deflection. The galvanometer and current stability did not result in more than a 2% change in calibration during any one magnetoresistance run. Since the magnetoresistance of the more highly concentrated samples was small while the zero-field

[10] E. W. Wallingford, thesis, University of Ottawa, Ottawa, Ontario (unpublished)).

[11] We would like to thank Dr. Salkovitz for his generosity in supplying these samples.

[12] E. I. Salkovitz, A. I. Schindler, and E. W. Kammar, Phys. Rev. 105, 887 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

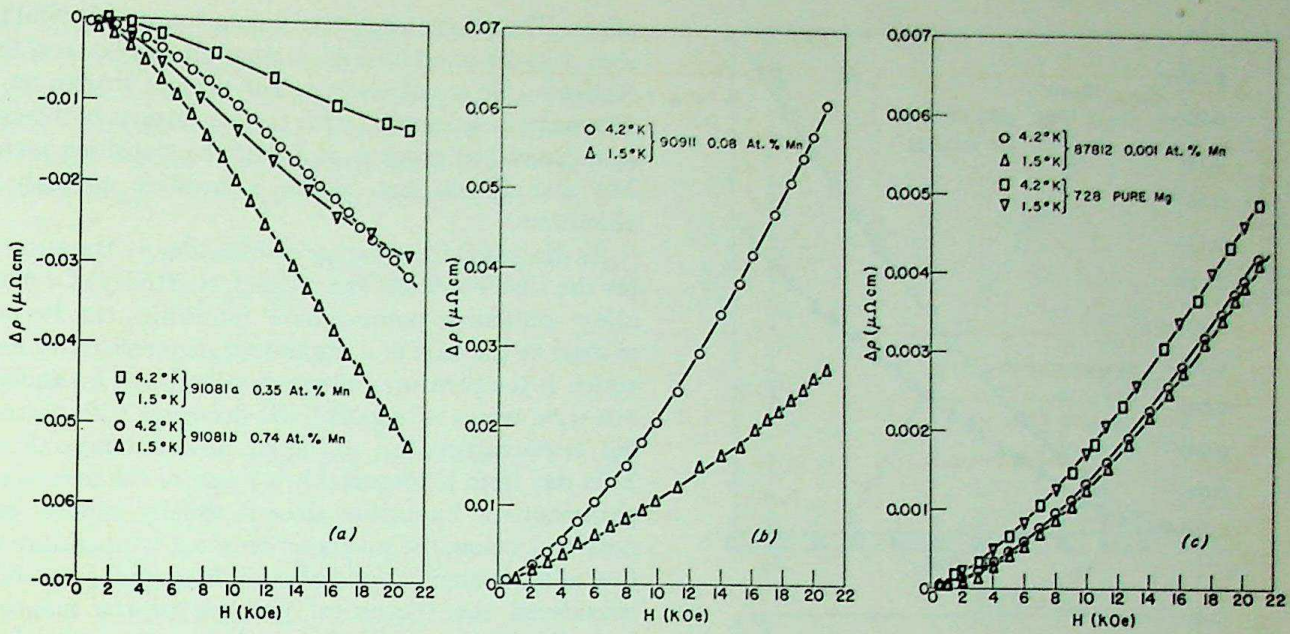


FIG. 2. $\Delta\rho$ as a function of the magnetic field strength for a number of Mg-Mn alloys: (a) for samples 91081(a) and (b); (b) for the sample 90911; (c) for samples 87812 and 728 (pure Mg).

resistance was comparatively high, the magnetoresistance measurements on these samples was accomplished by suppressing the zero-field sample voltage with the aid of a Leeds and Northrup Type K-2 potentiometer. A Harvey Wells Corporation electromagnet was used which produced fields of approximately 21 kOe in a gap of 1.25 in. with a stability of magnet current of one part in 10^5 at midcurrent values. All measurements were made at 4.2°K or lower by pumping on liquid helium. The vapor pressure of the bath was controlled by a Cartesian manostat.

3. EXPERIMENTAL RESULTS

Figure 1(a) shows the resistivity of some of the Mg-Mn alloys as a function of temperature in a constant magnetic field and it can be noticed that all of the samples exhibit anomalies in the temperature-dependent resistivity. At the maximum field available samples 87812 and 90911 exhibit a positive magnetoresistance, while both samples exhibit a resistance minimum. Samples 91081(a) and (b) exhibit an absolute decrease in the resistivity on the application of a magnetic field, the magnitude of the decrease increasing with magnetic field strength in the temperature region studied. These samples exhibit resistance maxima (and minima) in agreement with earlier reports.[7]

There is a marked anisotropy in the magnetoresistivity for Mg samples which is presumably due to preferred orientation induced by cold working the samples. The magnitude of the anisotropy can be seen for samples 728 and 87812 shown in Fig. 1(b). The maximum magnetoresistance values were used, and at this value very little angular dependence could be noted in the relative magnetoresistivity. For more concentrated samples, the anisotropy is so small that no significant difference appears for any chosen direction.

Figure 2 shows the magnetoresistance as a function of magnetic field strength for a number of specimens. As seen in Fig. 2(a) there is presumably sufficient magnetic scattering in the alloys to make the sign of the magnetoresistivity totally negative for samples 91081(a) and (b). Samples 90911 in Fig. 2(b) and sample 87812 in Fig. 2(c) show that $\Delta\rho$ is smaller at the lower temperature, while no such temperature dependence is observed for a pure magnesium sample exhibiting no anomalous behavior in the zero field resistivity. [See sample 728, Fig. 2(c).] It can therefore be suggested that there is a negative magnetoresistivity component in the more dilute Mg-Mn samples. Figures 3(a) and 3(b) show $\Delta\rho/\rho_{4.2°K}$ as a function of magnetic field strength for a sample of Mg-Cd and Mg-Al, respectively. As would be expected, $\Delta\rho/\rho_{4.2°K}$ increases with increase in magnetic field, and the magnitude of the change at fixed field decreases with impurity concentration increase (i.e., normal impurity scattering). Figures 4(a) and 4(b) show $\Delta\rho/\rho_T$ as a function of magnetic field for an Al-Fe and Al-Mn sample, where it can be seen that there is no temperature dependence to the magnetoresistance in these alloys.

4. DISCUSSION

The change in resistivity $\Delta\rho$ in a magnetic field H is given by the Kohler rule as

$$\Delta\rho/\rho_T = F(H\rho_\theta/\rho_T), \qquad (1)$$

where[13] ρ_θ is the resistivity at the Debye temperature θ

[13] As there is some uncertainty in the value of θ to be used in alloys, most workers approximate by using the value of resistivity at the ice point and we will follow the same approximation.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

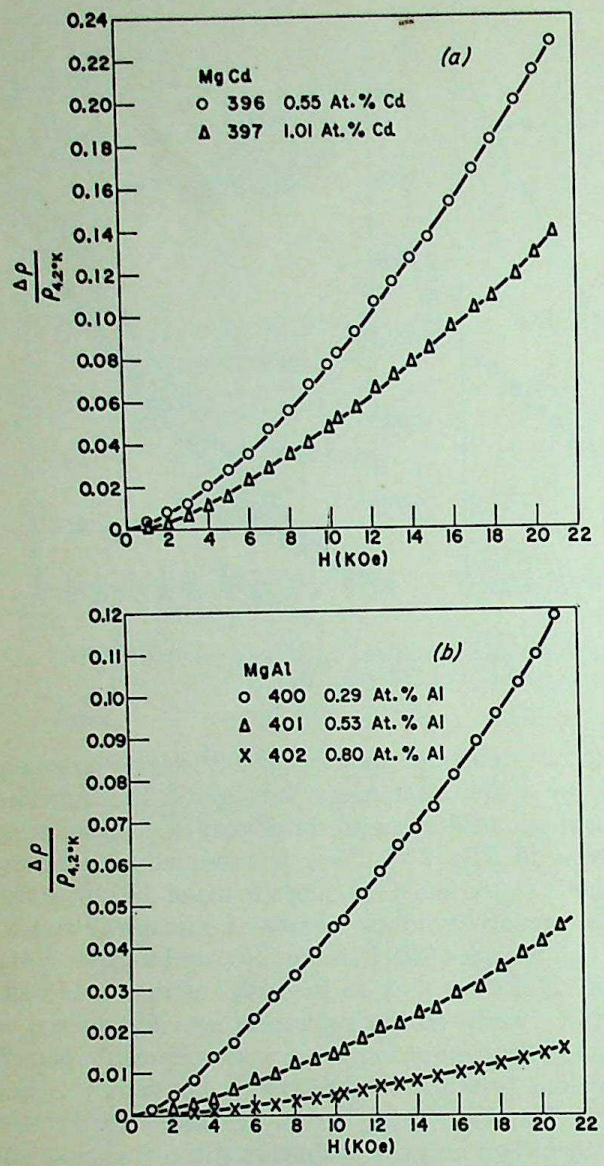


FIG. 3. $\Delta\rho/\rho_{4.2°K}$ as a function of magnetic field strength: (a) Mg-Cd and (b) Mg-Al alloys.

and ρ_T the resistivity at temperature T, the functional form of F being unspecified. Kohler's graph for some of the Mg and Al alloys is shown in Figs. 5 and 6, where it can be seen that all of the alloys presented obey the Kohler's rule. It can be noticed that the deviation from Kohler's rule is larger for the Mg-Al alloys than for the Mg-Cd alloys, and it could be speculated that this is a result of the difference in valence between solute and solvent in the case of the Mg-Al alloys. In Fig. 5 is the magnetoresistivity data at 4.2°K for the more dilute Mg-Mn alloys. It can be seen that Kohler's rule is obeyed which is in agreement with the previous conclusions of Webber,[5] Yntema,[14] and Thomas and Mendoza.[6] As we will see later this Kohler behavior is a result of the fact that although there is a negative component in the magnetoresistivity of these alloys it is much smaller than the normal positive magnetoresistivity of these alloys. The magnetoresistivity data for sample 90911 is shown as an insert in Fig. 5, where it can be seen that Kohler's rule is not obeyed. The present studies on Al alloys are in agreement with the conclusions of Thomas *et al.*,[6] and Hedgcock *et al.*,[8] that the transition metals Mn and Fe do not induce a resistive anomaly in aluminum.

As discussed in an earlier publication by the authors[3] for the Cu-Fe system, the magnetoresistivity of a dilute alloy containing paramagnetic impurities can be considered as the sum of a normal magnetoresistivity term which is temperature independent (ρ_n) and an anomalous term which is negative and decreases with decreasing temperature (ρ_s) at liquid helium temperatures. This last term is presumably a result of the presence of paramagnetic impurities since it usually appears only near and below the magnetic ordering temperature for the alloy system. Yosida[15] and Kasuya[16] have both considered the theoretical problem of the magnetic interaction between conduction electrons and localized magnetic ions where an *s-d* exchange interaction is assumed to produce a molecular field. Below a magnetic ordering temperature (usually antiferromagnetic in nature) theory predicts a monotonic decrease of resis-

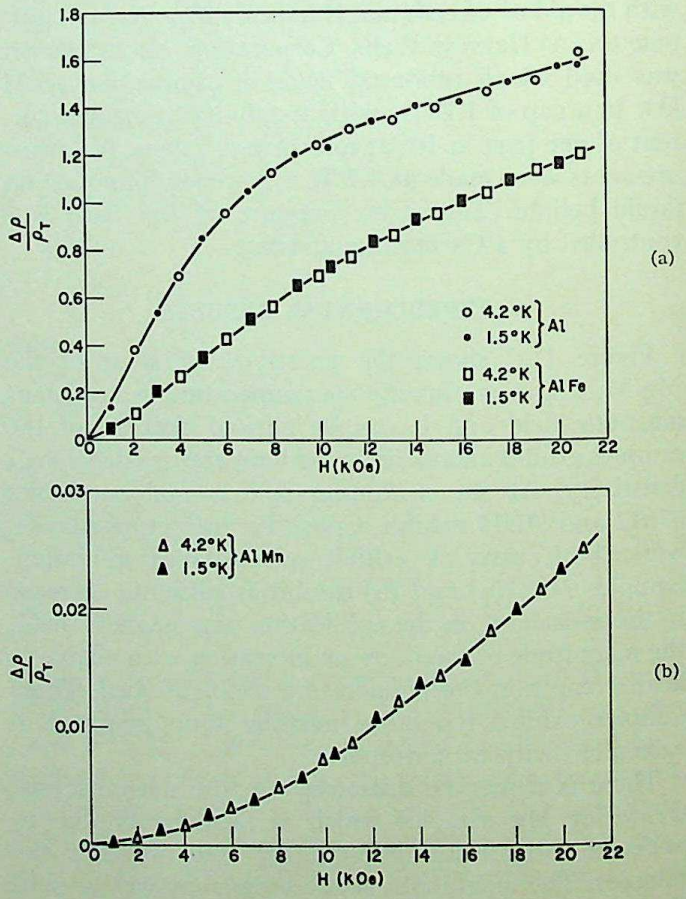


FIG. 4. $\Delta\rho/\rho_{4.2°K}$ as a function of the magnetic field strength at 4.2 and 1.5°K: (a) for pure Al and an Al-Fe alloy; (b) the Al-Mn alloy.

[14] G. B. Yntema, Phys. Rev. **91**, 1388 (1953).

[15] K. Yosida, Phys. Rev. **107**, 396 (1957).

[16] T. Kasuya, Progr. Theoret. Phys. (Kyoto) **22**, 227 (1959).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

tance as the temperature is lowered. Both theories predict a negative magnetoresistance which is proportional to the square of the magnetization at low magnetic field strengths. Neglecting a small term in Kasuya's expressions for the magnetoresistivity under various field and temperature conditions, we can write

$$H=0,\ T=0 \quad \rho_0=\frac{3\pi}{2}\frac{m^*}{\hbar\zeta}\frac{C}{Ne^2}[A^2+(g-1)^2j^2J^2], \tag{2}$$

$$H=0,\ T\geq T_N \quad \rho_{T_N}=\frac{3\pi}{2}\frac{m^*}{\hbar\zeta}\frac{C}{Ne^2}\times[A^2+(g-1)^2j(j+1)J^2], \tag{3}$$

$$H\neq 0,\ T\neq 0 \quad \rho=\frac{3\pi}{2}\frac{m^*}{\hbar\zeta}\frac{C}{Ne^2}\frac{[A^2+\langle j^2\rangle_{\rm av}J^2]^2}{A^2+\left(\langle j_z^2\rangle_{\rm av}-\bar{j}_z\frac{2xe^x}{(e^x-1)^2}\right)J^2}, \tag{4}$$

where m^* and e are the electron effective mass and charge, respectively, ζ is the Fermi energy, C is the concentration of paramagnetic ions, N is the number of atoms/cm^3, J is the s-d exchange interaction constant, A is the Coulomb interaction constant, $x=g\mu_BH/kT$, $\bar{j}_z$ is the average z component of the angular momentum j of the paramagnetic ion, and $\langle\bar{j}_z^2\rangle_{\rm av}$ is the average of the square of the angular momentum of the paramagnetic ion. For the method of finding these averages, the reader is referred to the original publication of Kasuya.[16] Using a simple molecular-field model, $H=H_0$ when $H_{\rm app}=0$, and if H_0 is zero, then it has the value $H_{\rm app}$ in the presence of an external field. Collings and Hedgcock[9] have shown through electron spin resonance and susceptibility studies that an antiferromagnetic transition occurs in the Mg-Mn system in the region of the resistance maximum. If one neglects the presence of the resistance minimum in these alloys and assumes the temperature of the maximum to be the Néel temperature,[17] we can use Eqs. (2) and (3) to calculate A and

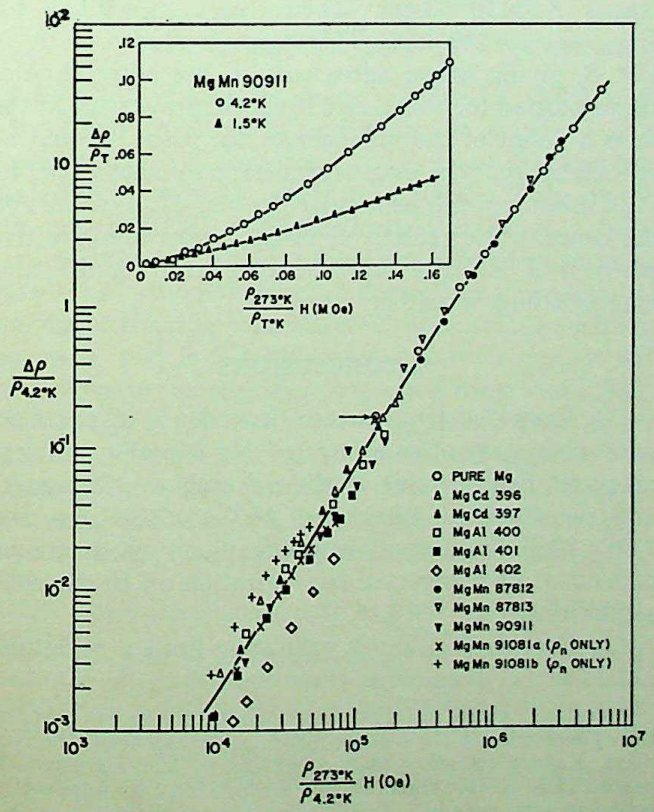


FIG. 5. $\Delta\rho/\rho_{4.2°K}$ as a function of $(\rho_{273°K}/\rho_{4.2°K})\cdot H$ for Mg samples. In insert is shown $\Delta\rho/\rho_T$ as a function of $(\rho_{273°K}/\rho_T)\cdot H$ for the sample 90911 at 4.2 and 1.5°K.

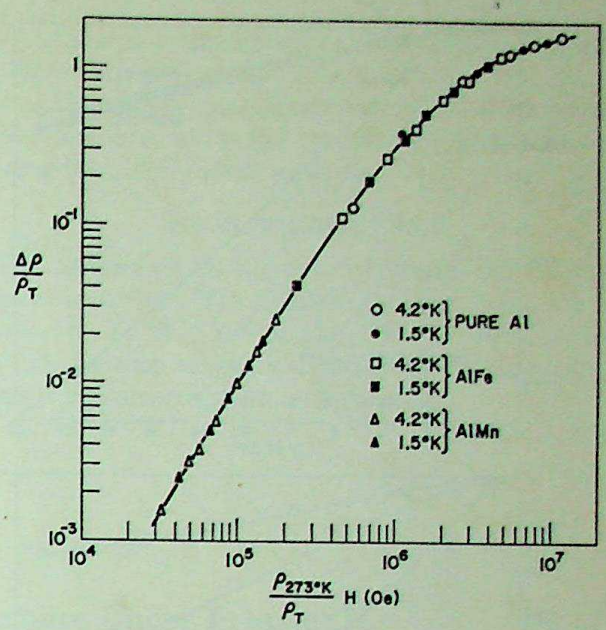


FIG. 6. $\Delta\rho/\rho_T$ as a function of $(\rho_{273°K}/\rho_T)\cdot H$ for Al samples.

TABLE II. J: the s-d exchange integral; A: Coulomb interaction constant; and A/J values of some MgMn alloys. The most reliable values for A and J are estimated from the data of Gaudet *et al.* (Ref. 7), where the resistivity was measured below 1°K.

Concentration (at.%)	J (10^{-12} erg)	A (10^{-12} erg)	A/J	
0.35	0.491	2.87	5.85	Present investigation
0.80	0.466	2.77	5.95	Wallingford's thesis
0.72	0.498	2.76	5.55	Wallingford's thesis
0.46	0.478	2.78	5.82	Wallingford's thesis
0.16	0.473	2.74	5.80	Gaudet *et al.*
average	0.48	2.78	5.80	

[17] Magnetic studies on these alloys (Ref. 9) have shown that only in the most concentrated sample is there any evidence for a Néel temperature but there are deviations from Curie's law in the region of the resistance anomaly for all of the Mg-Mn sample studies. Whether the temperature of the resistance maximum occurs at what one normally considers the Néel temperature is certainly still an open question theoretically. [See M. Baiyln, Westinghouse Research Report 029-B000-P1, 1961 (unpublished).]

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

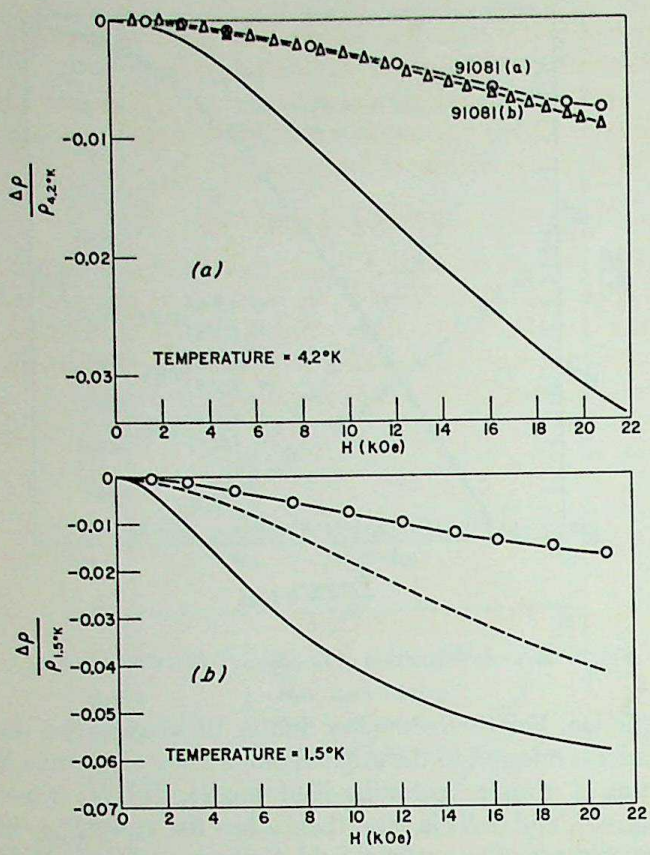


FIG. 7. (a) $\Delta\rho/\rho_{4.2°K}$ as a function of the magnetic field. ○: experimental results for the sample 91081(a). △: experimental results for the sample 91081(b), and ——— the calculated curve (ρ_s) based on expression (4). (b) $\Delta\rho/\rho_{1.5°K}$ as a function of the magnetic field. ○: experimental results for the sample 91081(a). ———: the calculated curve (ρ_s) based on expression (4). – – – –: the corrected curve (ρ) after allowance has been made for ρ_n.

J. This has been done for a number of alloys and the resulting values of A and J are listed in Table II. It can be seen that the values so derived for the various alloys are independent of concentration and result in an average value of 0.48×10^{-12} erg for J and 2.78×10^{-12} erg for A. The manganese ion is assumed to be in an S state with the value of the effective magneton number as derived from susceptibility studies ($S=2$). The free ion value of J derived for the Mn^{+} ion would yield a value of 0.35×10^{-12} erg, whereas a value of J for the Cu-Mn system derived from the experimental results is 0.67×10^{-12} erg.

Using the values of A and J derived from the zero-field results and expression (4) of Kasuya, it is possible to calculate the magnetoresistance as a function of magnetic field. Shown in Fig. 7(a) is the result of such a calculation for a sample 91081(a) and (b) at the temperature of 4.2°K. The difference between the calculated and experimental curves can be assumed due to the positive normal magnetoresistance component ρ_n. If the difference between the theoretical and experimental curves is in fact equal to ρ_n, then the difference should obey the Kohler rule. This difference is shown in Fig. 5 where as good a fit to the Kohler rule is found as for alloys containing nontransition metal impurities. Figure 7(b) shows the same calculation done at 1.5°K. The solid curve corresponds to that calculated from Eq. (4) and the dashed curve after allowance has been made for ρ_n, the normal impurity scattering discussed above. As can be seen, the agreement between experiment and theory at this temperature is not good. If one suggests a cooperative magnetic interaction and a resulting internal field, H_0 will shift the zero on the field axis but will not alter the magnitude of the magnetoresistance. Only by altering A/J could agreement with experiment for a particular magnetic field be found, but it is unlikely that these constants should be a function of temperature or magnetic field—certainly no theoretical treatment suggests that they should be.[18] It therefore seems that some essential feature is lacking in the Kasuya theory for the temperature dependence of the magnetoresistance in metal alloys containing transition metal impurities.

Although there is no adequate theory to explain the resistance minimum, it is interesting to apply Kasuya's expression for the magnetoresistivity in those alloys exhibiting a resistance minimum only [Eq. (4)]. For example, by assuming the same value of J and A as for the more concentrated alloys exhibiting a maximum, an estimate of $\Delta\rho$ at maximum magnetic field for sample 90911 yields $\Delta\rho_s=-2.4\times10^{-8}$ Ω-cm at 4.2°K. This can be compared with the experimentally observed value of $\Delta\rho=6.1\times10^{-8}$ Ω-cm. Therefore, the normal magnetoresistivity term, the difference of these two values, equals 8.5×10^{-8} Ω-cm. Then $\Delta\rho_n/\rho_{4.2°K}=0.152$ for $H\rho_{273}/\rho_{4.2°K}=149$ kOe. This point is shown as A in Fig. 5, giving better agreement to the initial Kohler curve plotted for this alloy. If this improvement in the fit is a result of the influence of $\Delta\rho_s$, then it could be speculated that the magnetoresistance of alloys exhibiting a minimum will be negative[19] when A/J is small and the concentration relatively large and becomes positive when A/J is large and the concentration of paramagnetic ions is smaller.

5. CONCLUSIONS

(1) From the temperature dependence of the electrical resistivity of concentrated Mg-Mn alloys in zero magnetic field a value of the s-d exchange integral J and the Coulomb integral Q of 0.48×10^{-12} erg and 2.78×10^{-12} erg, respectively, are found which are independent of manganese concentration up to the solid solubility limit.

(2) For Mg-Mn alloys exhibiting both a resistance

[18] Yosida pointed out that experimental agreement with his theory worsened as the temperature was lowered below the Néel point. Kasuya remarked on the discrepancy and suggested the magnetically ordered states should be represented by some sort of collective spin mode at sufficiently low temperatures.

[19] A total negative magnetoresistivity has been observed in Cu-Fe, Cu-Mn and Zn-Mn alloys exhibiting only a resistive minimum. (See Refs. 3 and 4.)

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

minimum and maximum the total magnetoresistance is negative. This magnetoresistance can be separated into a negative (ρ_s) and positive (ρ_n) component. The spin component ρ_s can be explained using the same J and A values as derived above at temperatures near the Néel temperature. The normal component (ρ_n) is found to obey Kohler's rule. The observed temperature dependence of the negative component (ρ_s) does not follow from the existing theory.

(3) For the more dilute Mg-Mn alloys exhibiting only a resistance minimum, there appears to be a negative magnetoresistivity component again requiring the same A and J values as above. The positive magnetoresistance component obeys Kohler's rule.

(4) Magnesium containing nontransition impurities follow Kohler's rule as do the alloys of aluminum containing transition metal impurities.

ACKNOWLEDGMENTS

The authors would like to acknowledge the technical assistance of Mrs. H. Hohnke and G. Mark, and to thank Dr. E. W. Collings, Dr. W. B. Muir, and D. Mathur for their suggestions and numerous discussions during the course of these experiments.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

F-Band Shape in the CsCl Structure Under Pressure*†

D. B. FITCHEN‡
Department of Physics, University of Illinois, Urbana, Illinois
(Received 24 January 1964)

Color centers have been formed by x irradiation of RbCl above the polymorphic transition pressure, and the shape of the new F band at 509 mμ in this CsCl-type structure has been examined under a pressure of 6200 atm at liquid-helium temperature. The band shows no evidence of the multiplet structure observed in the cesium halides, thus indicating that the lattice structure is not responsible for this effect. The F band in CsCl has also been formed and studied under pressure at helium temperature, and the triplet components show the usual relative intensity and the characteristic F-band shift. No evidence of K' bands was observed for these hydrostatic pressures.

INTRODUCTION

RECENT studies of color centers in CsCl and CsBr[1–3] revealed structure on the principal absorption band, in contrast to the usual F band in alkali halides with the NaCl-type lattice. This structure, which Rabin and Schulman[2] decomposed into three components at liquid-helium temperature, appears to be identical for centers formed by x irradiation or additive coloring, is unaffected by bleaching or heat treatment, and is not sensitive to sample purity. It therefore seems reasonable to associate these multiplets with a single electron-excess center, presumably an F center composed of an electron bound to a negative-ion vacancy. Rabin and Schulman[2] first proposed that if these components are all due to one center, then they must represent a splitting of the excited F state by a crystal-field effect peculiar to the CsCl-type lattice, or a splitting caused by the cesium ion itself.

The present work was undertaken to determine the influence of crystal structure by studying the F band formed in the pressure-induced simple cubic phase of a normally face-centered cubic alkali halide. RbCl transforms from the NaCl lattice structure to the CsCl lattice structure at about 5000–6000 atm. On the basis of a Mollwo relation between peak position and lattice constant for other CsCl-type crystals, one expects the F center to appear at about 520 mμ in the high-pressure form. Jacobs[4] looked for this peak in RbCl crystals x rayed or additively colored at atmospheric pressure prior to the structure transformation, but was unable to detect any new center. Drickamer *et al.*[5,6] found a new band, which they called the K' band, which replaced the F band in simple cubic phases at very high pressures and room temperature.

With the apparatus described below, it is possible to transform the crystal first and then form the CsCl-type centers by x raying under pressure at room or low temperatures. In particular it is possible to form

* This work was supported in part by the U. S. Atomic Energy Commission.

† Based on part of a thesis submitted to the University of Illinois in partial fulfillment of the requirements for the Ph.D. degree (1962). A preliminary report of this aspect of the work was given by D. Lazarus at the Conference on Crystal Lattice Defects, Kyoto, Japan, September 1962.

‡ Present address: Department of Physics, Cornell University, Ithaca, New York.

[1] P. Avakian and A. Smakula, Phys. Rev. **120**, 2007 (1960).
[2] H. Rabin and J. H. Schulman, Phys. Rev. Letters 4, 280 (1960); Phys. Rev. **125**, 1584 (1962).
[3] D. W. Lynch, Phys. Rev. **127**, 1537 (1962).
[4] I. S. Jacobs, Phys. Rev. **93**, 993 (1954).
[5] W. G. Maisch and H. G. Drickamer, Phys. Chem. Solids **5**, 328 (1958).
[6] R. A. Eppler and H. G. Drickamer, J. Chem. Phys. **32**, 1418 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and study the new F center at helium temperature to investigate whether it has the triplet structure observed in the cesium halides.

In addition, the F band was formed in CsCl under pressure at helium temperature to verify that the triplet components grow in the same ratio and exhibit the characteristic F-band pressure shift.

EXPERIMENTAL PROCEDURES

The four-window pressure cell and cryostat used in this experiment have been described in detail elsewhere.[7] Color centers were produced at high pressure and low temperature by irradiating the sample with 50-kV x rays through a beryllium window in the pressure cell. Optical absorption spectra were measured at right angles through sapphire windows, using a Beckman DU spectrophotometer. The sample was suitably masked to measure only a region of nearly uniform coloration, and the radiation-induced coloration was determined by subtracting the pre-irradiation spectrum from those measured at later times. Helium gas was used to apply hydrostatic pressure in the range 0–8000 atm, the pressure being determined with a manganin gauge except at temperatures where the helium solidified, when thermodynamic PVT data for solid helium were used to estimate[7] the pressure in the cell.

RESULTS AND DISCUSSION

a. RbCl

Two runs were made using samples 0.43 and 0.27 mm thick cleaved from a piece of "pure" Korth RbCl. Pressure was applied at room temperature at a rate of about 300 atm per min, and the transition pressure was detected by monitoring the optical transmission of the sample. At about 7300 atm, the transmission dropped abruptly to 0.17% in the first case and 1% in the second case. These values correspond to an optical density of roughly 80 cm^{-1}.

The crystallographic aspects of this type of structure transformation, which achieves a 15% reduction in volume,[8] have been discussed by Buerger[9] and applied to the phenomenon of opacity by Jacobs.[4] The transformation proceeds from a large number of nucleation centers giving rise to many mismatched grains. Light is scattered at the boundaries of these microscopic regions causing the observed opacity.

The transition in RbCl has been observed at 5700 atm by Bridgman,[8] 6000 atm by Pierce,[10] and 6500 atm by Jacobs, who also used the opacity as an indicator. The higher transition pressure seen in this case could be due to three causes. In another study of the polymorphic transformation in RbI with helium gas as a pressure fluid, Jacobs[11] found that the initial transformation for a virgin crystal occurred at about 10% higher pressures with the purely hydrostatic helium pressure medium than for liquid media. Apparently, the slight shear stresses due to viscosity of the liquid help to initiate the local shear involved in the transformation. A second reason might be hysteresis during the relatively rapid pressure rise, since Pierce found that the initial transformation required nearly 45 min to go to completion. However, no change in transmission was seen until the sudden abrupt and complete change. Finally, the crystals of less than "atomic-weight purity" usually have a higher transition pressure, but Pierce observed the transition at a lower pressure in Korth RbCl which had been doped with additional divalent impurity. Therefore, the first explanation seems the most probable here.

After the crystal had been transformed, it was cooled to helium temperature and x rayed under pressure. A preliminary run on the coloration of Korth RbCl at helium temperature below the transformation pressure had shown that the normal F band grows very slowly at about 590 mμ. In the present runs, the principal absorption grew slowly at about 510 mμ. Because of the very high background absorption, only the range of 420–620 mμ could be measured with any certainty, and readings outside this range fluctuated erratically. No systematic growth of any other major center was detected.

The shape of the new center was measured carefully twice in each run, and no indication of structure appeared, as shown in Fig. 1.

In the first run, a small peak was seen at 600 mμ, presumably due to NaCl-type F centers formed in portions of the crystal which had not completely transformed. When pressure was released slowly at nitrogen temperature, very marked hysteresis was observed and the new centers were unaffected to about 3500 atm, at which point they disappeared and the NaCl-type F center appeared with almost identical magnitude at 600 mμ. In the second run, the crystal was allowed to warm to room temperature under pressure and was then x rayed again. Instead of growth, the irradiation appeared to remove most of the residual centers which remained after warming. Similar hysteresis occurred when releasing pressure at room temperature. The optical density only began to decrease at about 4500 atm. At atmospheric pressure, the transmission had returned to about 20% of its initial value. The samples appeared milky and were highly strained but still intact.

The parameters of the new center as measured at various pressures and temperatures are listed in Table I.

While the designation of this new band as the F band is still tentative, the following evidence supports

[7] D. B. Fitchen, Rev. Sci. Instr. 34, 673 (1963).
[8] P. W. Bridgman, Z. Krist. 67, 363 (1928).
[9] M. J. Buerger, in *Phase Transformations in Solids* (John Wiley & Sons, Inc., New York, 1951).
[10] C. B. Pierce, Phys. Rev. 123, 744 (1961).
[11] R. B. Jacobs, Phys. Rev. 54, 468 (1938).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Parameters of the principal absorption band in transformed RbCl.

Temperature	Pressure (atm)	λ_m (mμ)	H (eV)
Helium	6400	509	0.22±0.01
Nitrogen	5230	517	0.25±0.01
Room	7370	514	0.35±0.03

it. The band appears to be the principal absorption center formed and has a shape and half-width similar to the F center in the low-pressure form. The peak of this new center shifts with temperature and pressure in the same manner as the F center, and its half-width increases with temperature in the same way. In the first run, it was possible to convert these centers directly to the normal F centers by returning to the low-pressure structure at nitrogen temperature. Finally, the new band appears at just the position predicted by a Mollwo-like relation for the CsCl structure, $\lambda_m = K(d)^n$, where d is the interionic distance. Using values of λ_m and d at low temperature for CsCl and CsBr,[2] and the value of d for the high-pressure form of RbCl from Bridgman's density measurements,[8] the new F center should appear at 514 mμ, in good agreement with experiment.

The lack of structure on this new F band in RbCl indicates that the triplet structure observed for CsCl and CsBr at liquid-helium temperature is not caused by a crystal-field effect indigenous to the CsCl-type lattice. Hughes and Rabin[12] have confirmed this result by their observation of clearly resolved peaks in the principal absorption band of CsF, the only cesium halide having the NaCl-type lattice. Thus the cause of the splitting seems to be the cesium ion itself.

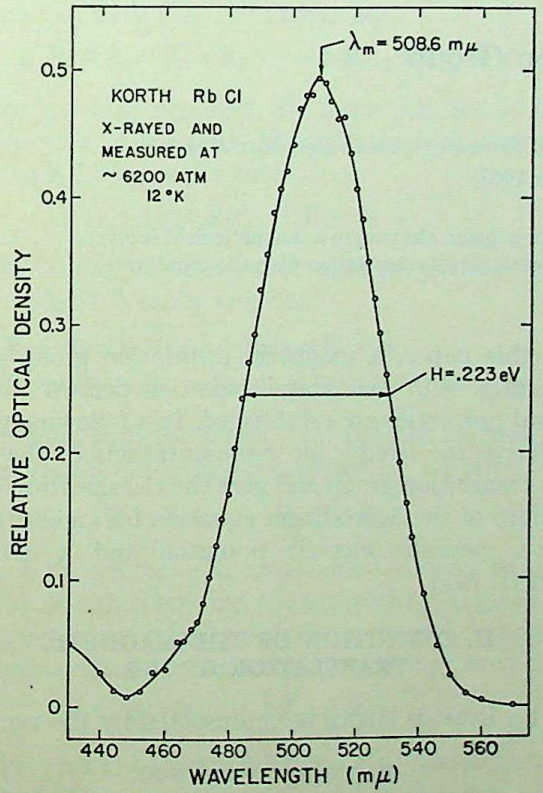


FIG. 1. Shape of the F band formed in the high-pressure CsCl phase of RbCl. The curve shows the difference in optical density measured before and after x irradiation for 220 min at this temperature and pressure for a sample 0.27 mm thick.

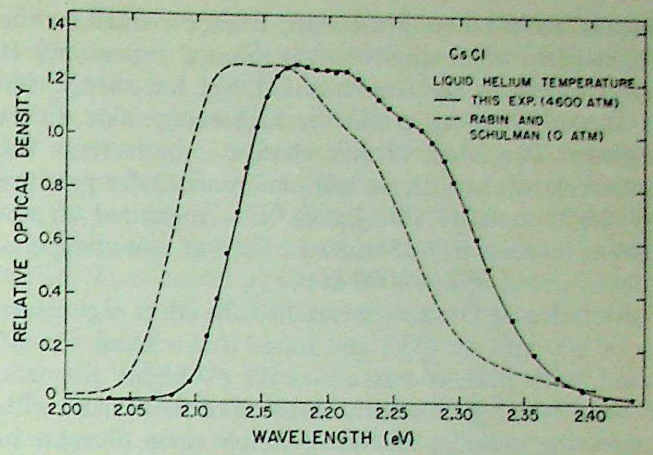


FIG. 2. Pressure shift of the F band in CsCl. The solid curve is for centers formed at this temperature and pressure after x irradiation for 15 min with a sample 0.84 mm thick. The comparison curve is taken from Ref. 2.

Speculation has centered on relativistic effects which would be appreciable only for the cesium halides. A calculation by Suffczynski[13] of the spin-orbit splitting for a localized excited state in a cesium halide gave a result orders of magnitude smaller than the observed effect. Recently, Knox[14] has proposed that a spin-orbit type splitting of the conduction band might be responsible. Calculations[15,16] of the other relativistic effects on band structure for heavy ions suggest that these might also contribute for cesium.

b. CsCl

A subsidiary experiment was performed to test whether the three components of the principal absorption band observed in CsCl at helium temperature would exhibit the characteristic pressure shift of the F band.

A freshly cut and polished platelet of Semi-Elements CsCl was x rayed and measured at approximately 4600 atm and 11°K. Figure 2 shows one of the measured spectra with a zero-pressure spectrum of Rabin and Schulman[2] normalized to the same peak height for comparison. The component structure is clearly visible and appears essentially unchanged after a pressure shift of $\Delta E/E \simeq 0.021$. (This shift is within 4% of

[12] F. Hughes and H. Rabin, Phys. Chem. Solids 24, 586 (1963).
[13] M. Suffczynski, J. Chem. Phys. 38, 1558 (1963).
[14] R. S. Knox, Phys. Rev. 133, A498 (1964).
[15] L. E. Johnson, J. B. Conklin, and G. W. Pratt, Jr., Phys. Rev. Letters 11, 538 (1963).
[16] F. Herman, C. D. Kuglin, K. E. Cuff, and R. L. Kortum, Phys. Rev. Letters 11, 541 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Jacobs' room-temperature shift[4] when corrected for the change in compressibility with thermal expansion.) If the two curves are superimposed, the low-energy side is almost identical, while the high-energy side differs slightly. The effect of this change is to increase the apparent total width at half-maximum under pressure by 3.5% to 0.017 eV. Jacobs[4] also measured an apparent increase in half-width in CsCl at room temperature of about 5% in 4000 atm.

Maisch and Drickamer[5] studied the effect of pressure to 50 000 atm on CsCl and found a new band, the K' band, with peak energy about 0.1 eV higher than the F band which grows at the expense of the F band with increasing pressure and is reversible upon decrease in pressure. It is thus possible that the high-energy component of the triplet is related to this K' band and is increasing in relative importance with pressure. However, no such K' band was seen in this range of hydrostatic pressures either for RbCl above or below its transition pressure, or for KBr, in which they observed a prominent K' band at considerably higher pressures. Either the K' band is not observable in the present range of pressures, or its appearance in Drickamer's experiments is due to shear and pressure inhomogeneities in his apparatus which are not present in our gas system.

SUMMARY

Using a novel technique a new absorption band has been formed by x irradiation of RbCl above its polymorphic transition pressure. The position, width, and temperature and pressure dependence of this band support its designation as the new F band in this CsCl-type phase. The band could be converted to the normal F band by reversing the transition at low temperatures. The shape of this new band has been examined under pressure at liquid-helium temperature, and shows no evidence of the multiplet structure observed in the cesium halides, thus indicating that the lattice structure is not responsible for this effect.

The pressure measurements on CsCl prove that for still another way of forming and observing the principal band, the triplet components all appear to be due to the F center.

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Professor David Lazarus, in whose laboratory this project was performed, for encouragement and advice throughout the work. He is also grateful to H. Rabin, W. D. Compton, and H. G. Drickamer for helpful discussions.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Magnetic Translation Group*

J. ZAK

National Magnet Laboratory,† Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 6 December 1963)

In this paper a group-theoretical approach to the problem of a Bloch electron in a magnetic field is given. A magnetic translation group is defined and its properties, in particular its connection with the usual translation group, are established.

I. INTRODUCTION

THE translation symmetry of the Hamiltonian for an electron in a periodic potential leads to the classification of the solutions of the corresponding problem by means of a wave vector $\mathbf{k}$ and to the possibility of introducing Bloch functions. When a constant magnetic field is also present, the Hamiltonian is no longer invariant under the translation group. However, operators may be defined which commute with the Hamiltonian of a Bloch electron in a magnetic field.[1,2]

In this paper, a magnetic translation group which commutes with the Hamiltonian is defined and its general properties are established. In a following paper, we derive the irreducible representations of the magnetic translation group and give the classification of the solutions of the Schrödinger equation for an electron in both a periodic electric potential and a constant magnetic field.

II. DEFINITION OF THE MAGNETIC TRANSLATION GROUP

Let a Bravais lattice be represented by the vectors

$$\mathbf{R}_n = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \quad (1)$$

(where $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are the unit cell vectors and n_1, n_2, n_3 are integers), each of which defines a point in the Bravais lattice. Let us define a path joining the origin O with the

* Supported by the U. S. Air Force Office of Scientific Research.

† Permanent address: Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel.

[1] P. G. Harper, Proc. Phys. Soc. (London) A68, 879 (1955); H. J. Fischbeck, Phys. Stat. Solidi 3, 1082 (1963).

[2] Recently, Brown [E. Brown, Bull. Am. Phys. Soc. 8, 257 (1963); Phys. Rev. 133, A1038 (1964)] has considered magnetic translation operators for constructing a ray representation of the usual translation group.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

point defined by $\mathbf{R}_n$. We start from point O and move in a direction of another lattice point, say, given by $\mathbf{R}_1$; from $\mathbf{R}_1$ we move in a direction, say, $\mathbf{R}_2$ to the point given by $\mathbf{R}_1+\mathbf{R}_2$ and so on until we achieve the point given by $\mathbf{R}_n$. Clearly, the point given by $\mathbf{R}_n$ can be achieved by different choices of vectors $\mathbf{R}_1$, $\mathbf{R}_2$ and so on. Let *one* of the possibilities be given by the vectors $\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_i$. We denote the described path by the symbol

$$|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i). \tag{2}$$

According to (2) we achieved the point of $\mathbf{R}_n$ by means of i steps. Also,

$$\mathbf{R}_n=\mathbf{R}_1+\mathbf{R}_2+\cdots+\mathbf{R}_i. \tag{2a}$$

In a similar way, we could choose another set of vectors satisfying (2a) and define a different path joining O with the point defined by $\mathbf{R}_n$.

Let us now define an operator which depends on both the vector $\mathbf{R}_n$ and the path joining the point O to the point defined by $\mathbf{R}_n$:

$$\begin{aligned}\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)&\\ =\exp\{(i/\hbar)\mathbf{R}_n\cdot[\mathbf{p}+(e/c)\mathbf{A}]\}&\\ \times\exp\{(i/2)[\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots&\\ +\mathbf{R}_{i-1}\times\mathbf{R}_i]\cdot\mathbf{h}\},&\end{aligned} \tag{3}$$

where $\mathbf{h}=e\mathbf{H}/\hbar c$, $\mathbf{H}$ is the magnetic field, $\mathbf{A}$ is the vector potential of the magnetic field, $\mathbf{p}$ is the momentum operator of the electron, and e is its charge. The order of the vectors in the vector products in (3) is given by their order in path (2). The expression

$$\tfrac{1}{2}[\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i]\cdot\mathbf{h} \tag{4}$$

has a very simple meaning. To show this, let us take the projection of the path (2) onto the plane perpendicular to $\mathbf{h}$. This gives a path

$$|\mathbf{R}_1{}^p,\mathbf{R}_2{}^p,\cdots,\mathbf{R}_i{}^p), \tag{5}$$

where $\mathbf{R}_k{}^p$ is the projection of $\mathbf{R}_k$ onto the plane perpendicular to $\mathbf{h}$. It is easily seen that

$$\begin{aligned}&\tfrac{1}{2}[\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i]\cdot\mathbf{h}\\ &=\tfrac{1}{2}[\mathbf{R}_1{}^p\times\mathbf{R}_2{}^p+\mathbf{R}_1{}^p\times\mathbf{R}_3{}^p+\cdots+\mathbf{R}_{i-1}{}^p\times\mathbf{R}_i{}^p]\cdot\mathbf{h}.\end{aligned} \tag{6}$$

The brackets with the factor $\frac{1}{2}$ on the right side of (6) give the area of the polygon enclosed by the vectors

$$\mathbf{R}_1{}^p,\ \mathbf{R}_2{}^p,\ \cdots,\ \mathbf{R}_i{}^p,\ -\mathbf{R}_n{}^p \tag{7}$$

(in which the vector $-\mathbf{R}_n{}^p$ was added in order to close the path). We thus see that the expression (4) gives the flux of the magnetic field (with the factor $e/\hbar c$) through the polygon which is obtained by projecting the vectors

$$\mathbf{R}_1,\ \mathbf{R}_2,\ \cdots,\ \mathbf{R}_i,\ -\mathbf{R}_n \tag{8}$$

onto the plane perpendicular to $\mathbf{H}$. Therefore, the elements (3) may be written in a different form

$$\begin{aligned}\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)=\exp\{(i/\hbar)\mathbf{R}_n\cdot[\mathbf{p}+(e/c)\mathbf{A}]\}&\\ \times\exp\{i\varphi(\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\},&\end{aligned} \tag{9}$$

where $\varphi(\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ is the flux of the magnetic field through the polygon (7) multiplied by $e/\hbar c$. It should be noted that the path (2), and therefore its projection onto the plane perpendicular to $\mathbf{H}$, can be very complicated. For example, from the point given by $\mathbf{R}_n$, which may be reached along the path (2), one may continue along the vector $-\mathbf{R}_n$ and then again turn back to point $\mathbf{R}_n$ by path (2). In such a case, the area in (6) will be doubled. The second exponent in the definition of the elements (3) or (9) thus expresses the curl nature of the magnetic field.

One may now show that the operators defined in (3) or (9) commute with the Hamiltonian of a Bloch electron in a magnetic field for a proper choice of the gauge. The Hamiltonian is

$$H=1/2m[\mathbf{p}-(e/c)\mathbf{A}]^2+V(\mathbf{r}), \tag{10}$$

m is the mass of the electron and $V(\mathbf{r})$ is the potential of the periodic field. It is easily shown that

$$\left[p_i+\frac{e}{c}A_i,\ p_k-\frac{e}{c}A_k\right]=i\hbar\frac{e}{c}\left[\frac{\partial A_k}{\partial x_i}+\frac{\partial A_i}{\partial x_k}\right]. \tag{11}$$

Therefore, if the gauge in (10) is chosen in a way that

$$\partial A_k/\partial x_i+\partial A_i/\partial x_k=0;\quad i,k=1,2,3, \tag{12}$$

the commutator (11) will vanish and the operators (3) will commute with the Hamiltonian (10). Relation (12) holds, for example, for a gauge

$$\mathbf{A}=\tfrac{1}{2}[\mathbf{H}\times\mathbf{r}]. \tag{13}$$

In order to show that the operators defined in (3) form a group, we shall check that a product of two such operators is again an operator of the same form. Let us take two operators

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\quad\text{and}\quad\tau(\mathbf{R}_n'|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k').$$

Their product is

$$\begin{aligned}&\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\tau(\mathbf{R}_n'|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k')\\ &=\exp\{(i/\hbar)(\mathbf{R}_n+\mathbf{R}_n')\cdot[\mathbf{p}+(e/c)\mathbf{A}]\}\\ &\times\exp\{(i/2)[\mathbf{R}_1\times\mathbf{R}_2+\cdots+\mathbf{R}_1\times\mathbf{R}_1'+\cdots\\ &+\mathbf{R}_{k-1}'\times\mathbf{R}_k']\cdot\mathbf{h}\},\end{aligned} \tag{14}$$

where the second exponent is defined by the path

$$|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i,\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k'). \tag{15}$$

In obtaining relation (14) we used the following expressions:

$$\left[\frac{i}{\hbar}\mathbf{R}_n\cdot\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right),\frac{i}{\hbar}\mathbf{R}_n'\cdot\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)\right]=i(\mathbf{R}_n\times\mathbf{R}_n')\cdot\mathbf{h} \tag{16}$$

(without any gauge limitations) and

$$e^Ae^B=e^{A+B+\frac{1}{2}[A,B]}$$

for A and B such that $[A,B]$ commutes with both

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

A and B. The product (14) can be written according to the definition (3) as follows:

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\tau(\mathbf{R}_n'|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k') = \tau(\mathbf{R}_n+\mathbf{R}_n'|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i,\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k'). \quad (17)$$

We now have a multiplication rule for the elements defined in (3): In order to multiply two elements [left side of (17)], we have to add the vectors $\mathbf{R}_n$ and $\mathbf{R}_n'$, and the paths $|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ and $|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k')$; the vector on which the product [right side of (17)] depends is $\mathbf{R}_n+\mathbf{R}_n'$ and the path is $|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i,\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k')$ [the latter is obtained by displacing the path $|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k')$ by the vector $\mathbf{R}_n$, thus obtaining a path joining point O with the point defined by the vector $\mathbf{R}_n+\mathbf{R}_n'$]. Thus, the product of two elements of the form (3) is an element of the same form. One sees that the elements defined in (3) form a group because one may also easily show that $\tau(-\mathbf{R}_n|-\mathbf{R}_i,-\mathbf{R}_{i-1},\cdots,-\mathbf{R}_1)$ is reciprocal to the element $\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$. We shall call this group the magnetic translation group (M.T.G.) and denote it by G.[2a] It should be noted that the operators

$$\exp\{(i/\hbar)\mathbf{R}_n\cdot(\mathbf{p}+(e/c)\mathbf{A})\} \quad (18)$$

do not form a group (as pointed out by Brown[2]), because a product of two such operators is not an operator of the form (18).

The group G defined in (3) is an infinite group for two reasons. First, there are an infinite number of vectors $\mathbf{R}_n$. Secondly, for any vector $\mathbf{R}_n$ an infinite number of different paths joining point O with $\mathbf{R}_n$ can be defined. We can, for instance, obtain from path P a different one P' by starting at and returning to any point belonging to the Bravais lattice on path P and adding a closed path to the path P at the mentioned point.

The structure of the group G becomes more apparent after its connection with the usual translation group, denoted by R, is established. To do this, let us define, for every vector $\mathbf{R}_n$, a set of elements $H(\mathbf{R}_n)$ to which all the elements $\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ of G belong. Each such set $H(\mathbf{R}_n)$ consists of an infinite number of elements. It is easy to see that we can look at the sets $H(\mathbf{R}_n)$ as at elements of a group. If we multiply an element of the set $H(\mathbf{R}_n)$ by an element of set $H(\mathbf{R}_n')$, we shall get an element of the set $H(\mathbf{R}_n+\mathbf{R}_n')$. $H(\mathbf{O})$ is the identity and $H(-\mathbf{R}_n)$ is reciprocal to $H(\mathbf{R}_n)$. This group, denoted by H, is an infinite commutative group, and is isomorphic to the translation group R. The isomorphism follows from the one-to-one correspondence between $\mathbf{R}_n$ and $H(\mathbf{R}_n)$. The group G is homomorphic to H,

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\rightarrow H(\mathbf{R}_n), \quad (19)$$

where the arrow shows the one-sided correspondence of the elements of G with those of H. From the homomorphism of G to H and the isomorphism of H and R, it follows that the magnetic translation group G is homomorphic to the usual translation group R. The last fact will be used to introduce the Born-von Karman boundary conditions for the representations of the magnetic translation group.

The group G has an invariant subgroup which is of great importance for constructing its irreducible representations. Consider the vectors r_m in the plane $\mathbf{a}_1$, $\mathbf{a}_3$ and construct elements of the group G:

$$\tau(\mathbf{r}_m|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_f') = \exp\{(i/\hbar)\mathbf{r}_m\cdot[\mathbf{p}+(e/c)\mathbf{A}]\} \times\exp\{(i/2)[\mathbf{R}_1'\times\mathbf{R}_2'+\mathbf{R}_1'\times\mathbf{R}_3'+\cdots +\mathbf{R}_{f-1}'\times\mathbf{R}_f']\cdot\mathbf{h}\}, \quad (20)$$

where

$$\mathbf{r}_m=m_1\mathbf{a}_1+m_3\mathbf{a}_3 \quad (21)$$

and $\mathbf{R}_1'$, $\mathbf{R}_2'$, $\cdots$, $\mathbf{R}_f'$ are arbitrary vectors of (1) except for the requirement that they form a path joining point O with the point defined by $\mathbf{r}_m$. The elements in (20) form a subgroup of G because a product of any two elements from (20) is also an element belonging to (20). Let us denote this subgroup by F. A similarity transformation by means of any element $\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ of G when applied to an element $\tau(\mathbf{r}_m|\mathbf{R}_1',\mathbf{R}_2',\cdots\mathbf{R}_f')$ of F gives the following result:

$$\tau^{-1}(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\tau(\mathbf{r}_m|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_f') \times\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i) = \tau(\mathbf{r}_m|-\mathbf{R}_i,-\mathbf{R}_{i-1},\cdots,-\mathbf{R}_1,\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_f',\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i). \quad (22)$$

The relation (22) shows us that subgroup F is an invariant subgroup of group G. It is clear that another invariant subgroup of G can be obtained by taking in (20) vectors $\mathbf{r}_m'$ in the plane $\mathbf{a}_2$, $\mathbf{a}_3$. This latter invariant subgroup is isomorphic to F. With respect to the invariant subgroup F, group G can be written as follows:

$$G=\tau(\mathbf{O}|\mathbf{O})F+\tau(\mathbf{a}_2|\mathbf{a}_2)F+\cdots +\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)F+\cdots. \quad (23)$$

It is of interest to note that when the magnetic field $\mathbf{H}$ is in the direction $\mathbf{a}_3$ the subgroup F is a commutative one, while the elements of F do not commute with those elements of G that do not belong to F.

So far, the groups have all been infinite. For constructing their irreducible representation, it is convenient to deal with finite groups. To do this, the Born-von Karman boundary conditions can be applied.

III. FINITE MAGNETIC TRANSLATION GROUP

The finiteness of the usual translation group is achieved by looking for its special representations, namely, by seeking representations which satisfy the

[2a] After the abstract of this paper appeared in Phys. Rev. Letters, Gerald A. Peterson pointed out (private communication) that in his Ph.D. thesis he has defined a closed set of magnetic translation operators and has called it magnetic translation group [G. A. Peterson, Ph.D. thesis, Cornell University, 1960 (unpublished)].

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Born-von Karman boundary conditions

$$D\{(\epsilon|\mathbf{a}_1)^N\}=D\{(\epsilon|\mathbf{a}_2)^N\}=D\{(\epsilon|\mathbf{a}_3)^N\}=D\{(\epsilon|0)\}\,, \quad (24)$$

where $(\epsilon|\mathbf{a}_1)$, $(\epsilon|\mathbf{a}_2)$, $(\epsilon|\mathbf{a}_3)$ are translations in the directions of the unit cell vectors (ϵ is the identity of the rotation group in the usual notation[3]), N is a large integer, and D denotes a representation of the translation group R. When seeking representations of R which satisfy the conditions (24), we can consider the group R as being a finite group $\bar{R}$ of order N^3, the elements of which are given by the direct product of the following three groups:

$$(\epsilon|\mathbf{a}_1)^i;\ (\epsilon|\mathbf{a}_2)^j;\ (\epsilon|\mathbf{a}_3)^k;\quad i,\,j,\,k=1,\,2,\,\cdots,\,N\,. \quad (25)$$

Similarly, in order to make G finite, we require similar boundary conditions to those imposed for the representations of the translation group (24). To do this, we use the fact that the translation group R is isomorphic to H and require the following conditions on the representations of H:

$$D\{H^N(\mathbf{a}_1)\}=D\{H^N(\mathbf{a}_2)\}$$
$$=D\{H^N(\mathbf{a}_3)\}=D\{H(\mathbf{O})\}\,. \quad (26)$$

With the conditions (26), the group H can be considered as being finite and given by the direct product of the following three groups:

$$\{H^i(\mathbf{a}_1)\};\ \{H^j(\mathbf{a}_2)\};\ \{H^k(\mathbf{a}_3)\};$$
$$i,\,j,\,k=1,\,2,\,\cdots,\,N\,. \quad (27)$$

Let us denote the last group by $\bar{H}$ and the corresponding M.T.G. by $\bar{G}$. It is clear that again $\bar{R}$ and $\bar{H}$ are isomorphic.

The requirements (26) mean that elements

$$\tau(Na_k|\mathbf{R}_1,\mathbf{R}_2,\cdots\mathbf{R}_f);\quad k=1,\,2,\,3\,, \quad (28)$$

of G can be considered as constant factors. From this it follows that the elements (28) commute with all the elements of G. It is easy to show that

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\tau(\mathbf{R}_n'|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k)$$
$$=\tau(\mathbf{R}_n'|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_k')\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$$
$$\times\exp\{i(\mathbf{R}_n\times\mathbf{R}_n')\cdot\mathbf{h}\}\,. \quad (29)$$

Hence, the elements (28) commute with all the elements of G if the exponential term in (29) equals 1 for $\mathbf{R}_n=N\mathbf{a}_k$ ($k=1,\,2,\,3$). We have therefore the following condition:

$$(N\mathbf{a}_k\times\mathbf{R}_n')\cdot\mathbf{h}=2\pi m;\quad k=1,\,2,\,3\,, \quad (30)$$

where m is an integer and $\mathbf{R}_n'$ is any vector from (1). Before discussing condition (30), let us derive another condition from the requirements (26). Consider two elements of the form (28) and take their product

$$\tau(N\mathbf{a}_i|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_f)\tau(N\mathbf{a}_k|\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_g')$$
$$=\tau(N\mathbf{a}_i+N\mathbf{a}_k|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_f,\mathbf{R}_1',\mathbf{R}_2',\cdots,\mathbf{R}_g')\,. \quad (31)$$

[3] G. F. Koster, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5.

But since the elements on the left side of (31) should behave as constant factors, the additional factor on the right side of (31),

$$\exp\{\tfrac{1}{2}(N\mathbf{a}_i\times N\mathbf{a}_k\cdot\mathbf{h}\}\,, \quad (32)$$

should be equal to 1. We have thus

$$(N\mathbf{a}_i\times N\mathbf{a}_k)\cdot\mathbf{h}=4\pi m;\quad i,\,k=1,\,2,\,3\,, \quad (33)$$

for odd N [for even N there is no additional condition to the one expressed by (30)]. Combining (30) and (33) we get for odd N the condition

$$(N\mathbf{a}_k\times\mathbf{R}_n')\cdot\mathbf{h}=4\pi m;\quad k=1,\,2,\,3\,, \quad (34)$$

which is stronger than the requirement (30). We now write the conditions (30) and (34) in a more convenient form, using the fact that (30) and (34) will be satisfied for any vector $\mathbf{R}_n'$ if they hold for the unit cell vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$:

$$N(\mathbf{a}_k\times\mathbf{a}_l)\cdot\mathbf{h}=2\pi n_{kl}\quad\text{(for even } N)\,, \quad (35a)$$

$$N(\mathbf{a}_k\times\mathbf{a}_l)\cdot\mathbf{h}=4\pi n_{kl}\quad\text{(for odd } N)\,, \quad (35b)$$

where n_{kl} is an integer and $k,\,l=1,\,2,\,3$. By using the definition of the reciprocal lattice vectors

$$K_i=(2\pi/V)\mathbf{a}_j\times\mathbf{a}_k\,, \quad (36)$$

where $i,\,j,\,k$ form the cycle (123) and V is the volume of this unit cell, we can write the conditions (35a), (35b) as follows:

$$N(V/2\pi)\mathbf{K}_m\cdot\mathbf{h}=2\pi m\quad\text{(for even } N)\,, \quad (37a)$$

$$N(V/2\pi)\mathbf{K}_m\cdot\mathbf{h}=4\pi m\quad\text{(for odd } N)\,. \quad (37b)$$

Here m is an integer. Equations (37a) and (37b) show that

$$\mathbf{h}=(2\pi/V)(\mathbf{R}_n/N)\quad\text{(for even } N)\,, \quad (38a)$$

$$\mathbf{h}=(4\pi/V)(\mathbf{R}_n/N)\quad\text{(for odd } N)\,. \quad (38b)$$

The requirements (38a), (38b) which follow from (26) are limitations on both the possible directions and values of the magnetic field. Let us choose the magnetic field in the direction $\mathbf{a}_3$ [this can always be done when conditions (38a), (38b) hold] and let $\mathbf{R}_n=n\mathbf{a}_3$; then

$$\mathbf{h}=(2\pi/V)(n/N)\mathbf{a}_3\quad\text{(for even } N)\,, \quad (39a)$$

$$\mathbf{h}=(4\pi/V)(n/N)\mathbf{a}_3\quad\text{(for odd } N)\,. \quad (39b)$$

For N sufficiently large the limitations (39a), (39b) on the magnetic field $\mathbf{H}$ are not essential. It should, however, be noted that when the crystal has dimensions of the order of 1 cm ($N\approx10^8$), the fields satisfying conditions (39a), (39b) differ by quanta of the order of 10 G (i.e., the lowest nonzero magnetic field satisfying the above conditions is of the order of 10 G). Unlike the boundary conditions for the usual translation group (which lead to no physical consequences), in the case of the M.T.G. we have restrictions on the magnetic field

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

even for crystals of the dimensions of 1 cm on which bulk experiments have been carried out.

We now show that the Born-von Karman conditions (26) on the representations of H, which led to requirements (39a), (39b), turn the group G into a finite one $\bar{G}$. By using (39a), (39b), we find that the elements of $\bar{G}$ are of the form

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)=\exp\{(i/\hbar)\mathbf{R}_n\cdot(\mathbf{p}+(e/c)\mathbf{A})\}$$
$$\times\begin{cases}\exp\{2\pi im(n/N)\} & \text{(for odd } N), \quad (40a)\\ \exp\{\pi im(n/N)\} & \text{(for even } N), \quad (40b)\end{cases}$$

where m is the coefficient by the product $\mathbf{a}_1\times\mathbf{a}_2$ in the expression

$$\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i. \quad (41)$$

From Eqs. (40a), (40b), we see that the number of elements of $\bar{G}$ is finite. In order to count the number of elements of $\bar{G}$, we treat the cases for even and odd N separately. For odd N, condition (39b) is imposed and the elements of $\bar{G}$ are given by (40a). The number of different values of the vector $\mathbf{R}_n$ is N^3. For each $\mathbf{R}_n$ there are different elements in (40a) arising from the second exponential factor. We first show that the number m in (40a) [and also in (40b)] can take all integer values. This follows from the fact that we can add to the path joining the point O with point $\mathbf{R}_n$ a closed path which will change (41) by an elementary area $\mathbf{a}_1\times\mathbf{a}_2$ and will change m in the expressions (40a), (40b) by unity. Let us now assume that p is the largest common factor of n and N (when there is no common factor, $p=1$), and let $N/p=N'$, $n/p=n'$. In order to count the different elements of (40a) for a given $\mathbf{R}_n$, let us check that

$$\exp\{2\pi im(n/N)\}\neq\exp\{2\pi im'(n/N)\} \quad (42)$$

for $m'\neq m$ (we exclude here those cases for which m' differs from m by the number N'). By assuming the equality sign in (42) to hold, we get

$$\exp\{2\pi i(m-m')(n/N)\}=1. \quad (43)$$

Relation (43) holds only when

$$(m-m')(n/N)=(m-m')(n'/N')=\text{integer}. \quad (44)$$

But since n' and N' have no common factor, the relation (44) is possible only for $m-m'\geq N'$, which means that m' should differ from m at least by N' in order to get the equality sign in (42). The second exponential in (40a) thus takes on N' different values for a given $\mathbf{R}_n$. Since $\mathbf{R}_n$ itself has N^3 different values, there are N^3N' different elements in $\bar{G}$ in the case of odd N. In a similar way, it is easy to show that there are $N^3(2N')$ different elements in $\bar{G}$ for even N.

It is of interest to note that when the magnetic field $\mathbf{H}$ is very strong (of the order of 10^{10} G), we can have the case that

$$\mathbf{h}=(4\pi/V)\mathbf{a}_3, \quad (45)$$

in which case the group $\bar{G}$ consists of N^3 commutative elements of the form

$$\exp\{(i/\hbar)\mathbf{R}_n\cdot(\mathbf{p}+(e/c)\mathbf{A})\} \quad (46)$$

and is isomorphic to the usual translation group.

IV. CONCLUSION

We have defined here a magnetic translation group G which commutes with the Hamiltonian for a Bloch electron in a magnetic field. In a following paper, we construct the irreducible representations of this group and give the classification of the solutions of Schrödinger's equation for an electron in both a periodic electric potential and a constant magnetic field.

ACKNOWLEDGMENTS

The author would like to thank the National Magnet Laboratory for its very kind hospitality and especially Dr. Arthur J. Freeman for helpful discussions and valuable suggestions in preparing the manuscript. The author would also like to thank Professor G. F. Koster for reading the manuscript and for his encouragement.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Magnetic Translation Group. II. Irreducible Representations

J. ZAK*

National Magnet Laboratory,† Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received 13 January 1964)

The physical irreducible representations of the magnetic translation group (M.T.G.) defined previously have been found. From these a set of solutions of Schrödinger's equation for a Bloch electron in a magnetic field has been constructed. In general the M.T.G. is non-Abelian. However, when the magnetic flux through areas enclosed by any vectors of the Bravais lattice become multiples of an elementary "fluxon" hc/e, the M.T.G. becomes isomorphic to the usual translation group.

I. INTRODUCTION

IN a previous paper[1] a magnetic translation group was defined and its general properties were established. The elements of this group, which depend on both a point of the Bravais lattice and a path joining the origin of the lattice with this point, are given by

$$\tau(\mathbf{R}|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)=\exp\{(i/\hbar)\mathbf{R}_n\cdot[\mathbf{p}+(e/c)\mathbf{A}]\}\times\exp\{2\pi i(n/bN)m\}, \quad (1)$$

where $\mathbf{R}_n$ is a vector of the Bravais lattice, $\mathbf{A}$ is the vector potential of the magnetic field $\mathbf{H}$, m is the coefficient by the product $\mathbf{a}_1\times\mathbf{a}_2$ of the unit cell vectors $\mathbf{a}_1$, $\mathbf{a}_2$ in the sum of ordered products

$$\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots+\mathbf{R}_1\times\mathbf{R}_i+\mathbf{R}_2\times\mathbf{R}_3+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i, \quad (2)$$

n, N, and b appear from the Born-von Karman conditions which turn the magnetic translation group into a finite one and impose the following condition on the magnetic field:

$$\frac{e\mathbf{H}}{\hbar c}=\frac{4\pi}{V}\frac{\mathbf{R}_n}{bN}=\frac{4\pi}{V}\frac{n}{bN}\mathbf{a}_3, \quad (3)$$

V is the volume of the unit cell, $V=\mathbf{a}_1\times\mathbf{a}_2\cdot\mathbf{a}_3$, and b takes the values 1 or 2 for N odd or even, respectively.

The finite magnetic translation group $\bar{G}$ is of the order $b(N/p)N^3$, where p is the largest common factor of N and n. For constructing the irreducible representations of $\bar{G}$ we use the fact that it has an invariant subgroup $\bar{F}$ consisting of elements

$$\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)=\exp\left\{\frac{i}{\hbar}\mathbf{r}_n\cdot\left(\mathbf{p}+\frac{e}{c}\mathbf{A}\right)\right\}\times\exp\left\{2\pi i\frac{n}{bN}m\right\}, \quad (4)$$

where

$$\mathbf{r}_n=n_1\mathbf{a}_1+n_3\mathbf{a}_3. \quad (5)$$

The group $\bar{G}$ can therefore be written as

$$\bar{G}=\tau(\mathbf{a}_2|\mathbf{a}_2)\bar{F}+\cdots+\tau(N\mathbf{a}_2|N\mathbf{a}_2)\bar{F}. \quad (6)$$

The invariant subgroup $\bar{F}$ is a commutative group and its irreducible representations are well known. In order to find the irreducible representations of $\bar{G}$ we use the method of Frobenius.[2,3] This method enables one to find all the irreducible representations of a group from the irreducible representations of its invariant subgroup.

In Sec. II the physical irreducible representations of the M.T.G. are constructed. They turn out to be the same in the case of the finite M.T.G. as the ray representations of the usual translation group that are given by Brown.[4] This result is not surprising because in the usual method for constructing ray representations of a finite group[5,6] one augments the group, finds the vector representations of the augmented group, and then chooses those representations which correspond to ray representations of the original group. It appears, however, more natural to define a group which commutes with the Hamiltonian for a Bloch electron in a magnetic field than to treat ray representations of a group (the usual translation group) which does *not* commute with this Hamiltonian. Moreover, by finding all the physical representations of the M.T.G., we are sure that we have all the possible symmetry types for the wave functions.[7]

For constructing the irreducible representations of the M.T.G., we first start with the finite group obtained by applying the boundary conditions. From the method of construction, it is possible to find a complete set of symmetry adapted functions for the problem under discussion.[8] These functions are presented in Sec. III. As shown in Ref. 1, the boundary conditions on the magnetic translation group lead to physical restrictions on the magnitude of the magnetic field. In Sec. IV we

---

* Permanent address: Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel.

† Supported by the U. S. Air Force Office of Scientific Research.

[1] J. Zak, Phys. Rev. 134, A1602 (1964).

[2] F. Seitz, Ann. Math. 37, 17 (1936).

[3] J. Zak, J. Math. Phys. 1, 165 (1960).

[4] E. Brown, Bull. Am. Phys. Soc. 8, 257 (1963); Phys. Rev. 133, A1038 (1964).

[5] I. Schur, J. Reihne Aug. Math. 127, 20 (1904); 132, 85 (1907).

[6] W. Döring, Z. Naturforsch. 14, 343 (1959).

[7] The statement used in Ref. 4 for showing the completeness of the irreducible representations, namely, that the sum of squares of the dimensions of nonequivalent irreducible ray representations equals the order of the group is, in general, not true. There are, for example, finite groups that have no ray representations at all.

[8] An alternative method for constructing symmetry adapted functions is discussed in Ref. 4.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

remove these restrictions by dealing with an infinite M.T.G. Finally, in the same section, the significance of the elementary fluxon hc/e for the magnetic translation group is discussed.

II. CONSTRUCTION OF THE IRREDUCIBLE REPRESENTATIONS OF G

The invariant subgroup $\bar{F}$ of $\bar{G}$ is a commutative group and its irreducible representations are

$$D^{(s,k_1,m_3)}\{\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\} = \exp\left\{2\pi i\left[\frac{n}{bN}ms+\frac{k_1n_1}{N}+\frac{m_3n_3}{N}\right]\right\}, \quad (7)$$

where s takes the values from 1 to bN/p, and k_1, m_3 the values from 1 to N. Relation (7) thus defines $(bN/p)N^2$ one-dimensional representations of $\bar{F}$. These are all the irreducible representations of $\bar{F}$ because the order of the latter is just $(bN/p)N^2$. However, some of the representations (7) are nonphysical as may be seen by setting $\mathbf{r}_n=\mathbf{o}$. The element $\tau(\mathbf{o}|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ is a constant factor which will be represented by different constants if we take different values of s in (7). When the operator $\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)$ itself operates in the functional space, the only physical representations in (7) will be those for $s=1$. We therefore consider the representations of $\bar{F}$ only for $s=1$ and from these we construct the irreducible representations of $\bar{G}$.

Let us first consider the case $p=1$ (i.e., no common factor for N and n) and denote the vector that spans representation (7) (for $s=1$) by $|k_1, m_3\rangle$. Then

$$D^{(k_1,m_3)}\{\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\}|k_1,m_3\rangle = \exp\left\{2\pi i\left[\frac{n}{bN}m+\frac{2nk_1n_1}{bN}+\frac{m_3n_3}{N}\right]\right\}|k_1,m_3\rangle, \quad (8)$$

where a factor $2n/b$ was introduced for convenience in front of k_1n_1/N. (This is equivalent to an alternative numbering of the eigenvectors $|k_1,m_3\rangle$.) We now perform a similarity transformation:

$$\tau(-n_2\mathbf{a}_2|-n_2\mathbf{a}_2)\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2) = \tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\exp\{4\pi i(n/bN)n_1n_2\}. \quad (9)$$

It follows from (9) that

$$D\{\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i\}D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k_1,m_3\rangle = \exp\left\{2\pi i\left[\frac{n}{bN}m+\frac{2n(k_1+n_2)n_1}{bN}+\frac{m_3n_3}{N}\right]\right\} \times D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k_1,m_3\rangle. \quad (10)$$

It is seen that the vector $D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k_1,m_3\rangle$ is an eigenvector of $D\{\tau(\mathbf{r}_n|\mathbf{R}_1,\mathbf{R}_2,\cdots,\mathbf{R}_i)\}$ corresponding to the eigenvalues k_1+n_2, m_3 (D now denotes a representation of $\bar{G}$). Since we assumed that N and n have no common factor we get[2] N vectors $D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k_1,m_3\rangle$ for $n_2=1, 2, \cdots, N$:

$$D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k_1,m_3\rangle = |k_1+n_2, m_3\rangle. \quad (11)$$

Let us show that relations (10) and (11) for each m_3 define an irreducible representation of $\bar{G}$. According to (6) each operator $\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)$ can be written as follows:

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i) = \tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\tau(\mathbf{r}_n|-n_2\mathbf{a}_2,\mathbf{R}_1,\cdots,\mathbf{R}_i), \quad (12)$$

where $\mathbf{R}_n=\mathbf{r}_n+n_2\mathbf{a}_2$, $\mathbf{r}_n=n_1\mathbf{a}_1+n_3\mathbf{a}_3$. Therefore,

$$D\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\}|k_1,m_3\rangle = \exp\left\{2\pi i\left[\frac{n}{bN}g+\frac{2nk_1n_1}{bN}+\frac{m_3n_3}{N}\right]\right\}|k_1+n_2, m_3\rangle, \quad (13)$$

where g is the coefficient of the product $\mathbf{a}_1\times\mathbf{a}_2$ in

$$-n_2\mathbf{a}_2\times\mathbf{R}_n+\mathbf{R}_1\times\mathbf{R}_2+\mathbf{R}_1\times\mathbf{R}_3+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i. \quad (14)$$

From (13) it follows that

$$D_{k'k}{}^{m_3}\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\} = \langle m_3k'|\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)|k,m_3\rangle = \delta_{k',k+n_2}\exp\left\{2\pi i\left[\frac{n}{bN}g+\frac{2nkn_1}{bN}+\frac{m_3n_3}{N}\right]\right\}, \quad (15)$$

where $D_{k'k}{}^{m_3}$ denotes a matrix element $k'k$ of the operator $\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)$; k' and k take values from 1 to N, and $k+n_2$ is taken modulo N.

We next show that for each m_3 the matrices D^{m_3} form a representation of $\bar{G}$. To do this we first calculate the product of two matrices (15) corresponding to two elements $\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)$ and $\tau(\mathbf{R}_n'|\mathbf{R}_1',\cdots,\mathbf{R}_i')$ and then the matrix which corresponds to the product $\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\tau(\mathbf{R}_n'|\mathbf{R}_1',\cdots,\mathbf{R}_i')$. The two results must be the same if (15) is to be a representation of $\bar{G}$. From (15) we have

$$\sum_{\alpha=1}^{N} D_{k'\alpha}{}^{m_3}\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\}D_{\alpha k}{}^{m_3}\{\tau(\mathbf{R}_n'|\mathbf{R}_1',\cdots,\mathbf{R}_i')\} = \delta_{k',k+n_2+n_2'}\exp\left\{2\pi i\left[\frac{n}{bN}(g+g')+\frac{2nk(n_1+n_1')}{bN}+\frac{2nn_1n_2'}{bN}+\frac{m_3(n_3+n_3')}{N}\right]\right\}. \quad (16)$$

On the other hand, the matrix corresponding to the product of the two elements is

$$D_{k'k}{}^{m_3}\{\tau(\mathbf{R}_n+\mathbf{R}_n'|\mathbf{R}_1,\cdots,\mathbf{R}_i,\cdots,\mathbf{R}_i')\} = \delta_{k',k+n_2+n_2'}\exp\left\{2\pi i\left[\frac{n}{bN}g''+\frac{2nk(n_1+n_1')}{bN}+\frac{m_3(n_3+n_3')}{N}\right]\right\}, \quad (17)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where g'' is the coefficient by the product $\mathbf{a}_1\times\mathbf{a}_2$ in the expression

$$-(n_2+n_2')\mathbf{a}_2\times(\mathbf{R}_n+\mathbf{R}_n')+\mathbf{R}_1\times\mathbf{R}_2+\cdots+\mathbf{R}_{i-1}\times\mathbf{R}_i+\mathbf{R}_1'\times\mathbf{R}_2'+\cdots+\mathbf{R}_{i-1}'\times\mathbf{R}_i'. \quad (18)$$

It is easily shown that

$$g''=g+g'+2n_1n_2'. \quad (19)$$

It therefore follows that expressions (16) and (17) are equal and thus that the matrices (15) for each m_3 form a representation of $\bar{G}$. These representations are irreducible because all the vectors $|km_3\rangle$ for a given m_3 and different $k(k=1, 2; \cdots, N)$ form a closed subspace. An alternative way to check the irreducibility of (15) is to calculate the character for a given m_3,

$$\chi^{m_3}\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\}=\sum_{k=1}^{N}D_{kk}^{m_3}=N_{\delta n_1 0\delta n_2 0}\exp\left\{2\pi i\left[\frac{n}{bN}g+\frac{m_3n_3}{N}\right]\right\}, \quad (20)$$

and to find the sum

$$\sum_{\bar{G}}|\chi^{m_3}\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\}|^2=N^4. \quad (21)$$

Result (21) shows that the representation (15) with the character (20) is irreducible [for reducible representations the sum (21) is greater than the number of elements in $\bar{G}$]. Note that the way of constructing irreducible representations here corresponds to the method of finding conjugate representations of the invariant subgroup $\bar{F}$ by means of the elements of $\bar{G}$ that do not belong[3] to $\bar{F}$. In the case when N and n have no common factor all the conjugate representations are nonequivalent and the representation of $\bar{G}$ which they induce is thus irreducible.

According to the method of Froebenius, relation (15) gives us all the irreducible representations of the group[2,3] corresponding to $s=1$. Since we are interested only in the physical representations of $\bar{G}$ these are all the representations of $\bar{G}$.

Let us now treat the case when the largest common factor of N and n is p. If we start with a vector $|k_1,m_3\rangle$ [see (8)] and apply all the operators $\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)$ we shall get in general only $N'=N/p$ independent vectors $|k,m_3\rangle$ for $k=1, 2, \cdots, N'$. According to Ref. 2 it again follows that we may assume

$$D\{\tau(n_2\mathbf{a}_2|n_2\mathbf{a}_2)\}|k,m_3\rangle=\exp\{2\pi i(m_2n_2/N)\}|k+n_2, m_3\rangle \quad (22)$$

for $n_2=1, 2, \cdots, N'$ and $m_2=0, 1, \cdots, p-1$. The representations induced by the vectors (22) can be written as

$$D_{k'k}^{(m_1m_2m_3)}\{\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\}=\delta_{k',k+n_2}\exp\left\{2\pi i\left[\frac{m_1n_1}{N}+\frac{m_2n_2}{N}+\frac{n}{bN}g+\frac{2nkn_1}{bN}+\frac{m_3n_3}{N}\right]\right\}, \quad (23)$$

where the values of m_1, m_2 range from 0 to $p-1$; k, k' assume values from 1 to N' ($k+n_2$ is defined modulo N'), and m_3 ranges from 1 to N. Relation (23) defines p^2N irreducible and nonequivalent $N'\times N'$ representations of the group $\bar{G}$ and again gives us all the physical representations of $\bar{G}$. Note that in this case when N and n have a common factor the conjugate representations of $\bar{F}$, constructed by means of elements of $\bar{G}$ that do not belong to $\bar{F}$, will in general no longer be nonequivalent. The representations of $\bar{G}$ induced by the conjugate representations of $\bar{F}$ will be reducible; the irreducible parts are given in (23). In the case when $p=1$, representations (23) go over into (15). Hence, relation (23) gives all the irreducible physical representations of $\bar{G}$.

III. SYMMETRY ADAPTED FUNCTIONS

We now use the irreducible representations of the magnetic translation group $\bar{G}$ to construct symmetry adapted solutions of Schrödinger's equation for a Bloch electron in a magnetic field. By using the gauge

$$\mathbf{A}=\tfrac{1}{2}[\mathbf{H}\times\mathbf{r}], \quad (24)$$

we can write the elements (1) of the magnetic translation group as

$$\tau(\mathbf{R}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)=\exp\left\{\frac{i}{\hbar}\mathbf{R}_n\cdot\mathbf{p}+i\frac{n}{bN}(n_2\mathbf{K}_1-n_1\mathbf{K}_2)\cdot\mathbf{r}\right\}\times\exp\left\{2\pi i\frac{n}{bN}m\right\}, \quad (25)$$

where $\mathbf{K}_1$ and $\mathbf{K}_2$ are vectors of the unit cell of the reciprocal lattice. Let

$$\psi(\mathbf{r})=\exp\{i\mathbf{k}\cdot\mathbf{r}\}u(\mathbf{r}), \quad (26)$$

with

$$\mathbf{k}=\frac{1}{N}\left[\left(m_1+\frac{2nk_1}{b}\right)\mathbf{K}_1+m_3\mathbf{K}_2\right],$$

$$m_1=0, 1, \cdots, p-1; \quad k_1=1, \cdots, N'; \quad (27)$$

and

$$\tau(\mathbf{r}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)u(\mathbf{r})=\exp\{2\pi i(n/bN)g\}u(\mathbf{r}).$$

The last condition on $u(\mathbf{r})$ means that

$$u(\mathbf{r}+\mathbf{r}_n)=\exp\left\{i\frac{n}{bN}\mathbf{n}_1\mathbf{K}_2\cdot\mathbf{r}\right\}u(\mathbf{r}), \quad (28)$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

in which case

$$\tau(\mathbf{r}_n|\mathbf{R}_1,\cdots,\mathbf{R}_i)\psi(\mathbf{r}) = \exp\left\{2\pi i\left[\frac{n}{bN}g+\frac{m_1n_1}{N}+\frac{2nk_1n_1}{bN}+\frac{m_3n_3}{N}\right]\right\}\psi(\mathbf{r}). \quad (29)$$

In order to satisfy condition (28) and the boundary conditions, we choose $u(r)$ as follows:

$$u(\mathbf{r}) = \exp\left\{\frac{i}{2\pi}\frac{n}{bN}\mathbf{r}\cdot\mathbf{K}_1\mathbf{r}\cdot\mathbf{K}_2\right\}w(\mathbf{r}), \quad (30)$$

with

$$w(\mathbf{r}+\mathbf{r}_n) = w(\mathbf{r}), \qquad w(\mathbf{r}+N'\mathbf{a}_2) = \exp\{-iN'\mathbf{a}_2\times\mathbf{h}\cdot\mathbf{r}\}w(\mathbf{r}). \quad (31)$$

The function (26) will therefore be

$$\psi(\mathbf{r}) = \exp\{i\mathbf{k}\cdot\mathbf{r}\}\exp\left\{\frac{i}{2\pi}\frac{n}{bN}\mathbf{r}\cdot\mathbf{K}_1\mathbf{r}\cdot\mathbf{K}_2\right\}w(\mathbf{r}). \quad (32)$$

By using Eq. (32) and the relation (22) and by reintroducing the magnetic field according to Eq. (3) we get the functions

$$\psi_q^{(l,\mathbf{k})}(\mathbf{r}) = \exp\{i\mathbf{k}\cdot\mathbf{r}\} \times\exp\{i[q+(1/4\pi)\mathbf{K}_2\cdot\mathbf{r}](\mathbf{a}_2\times\mathbf{h})\cdot\mathbf{r}\}w_{lk}(\mathbf{r}+q\mathbf{a}_2), \quad (33)$$

where $\mathbf{k}=(1/N)(m_1\mathbf{K}_1+m_2\mathbf{K}_2+m_3\mathbf{K}_3)$, m_1, m_2 are integers which take values from 0 to $p-1$ (with p again the common factor of N and n), m_3 is an integer whose values range from 1 to N and the integer q has values from 0 to $N'-1$. The functions $w_{lk}(\mathbf{r})$ satisfy conditions (31). The meaning of l will be given later. For fixed $\mathbf{k}$ and l there are $N'(=N/p)$ functions in (33) which form a basis for an $N'\times N'$ irreducible representation (23) of the magnetic translation group $\bar{G}$. For different $\mathbf{k}$ we get nonequivalent representations of $\bar{G}$. The functions in (33) are thus the symmetry adapted functions for all the irreducible representations of $\bar{G}$. The physical picture connected with this description is the following. All the functions (33) with fixed l and $\mathbf{k}$ (and varying q) belong to the same energy. When $\mathbf{k}$ also varies the energy will be in general varying and one obtains an energy band for each l. The number l was thus introduced to distinguish between different energy bands in the same way in which it is done in the case of the usual translation group.

From the fact that the function (33) form bases for irreducible unitary representations (23) of the magnetic translation group, it follows that functions (33) are orthogonal for different $\mathbf{k}$ and q. [Note that the functions $w_{lk}(\mathbf{r})$ also depend on the vector $\mathbf{k}$. This dependence is obtained by substituting functions (33) into Schrödinger's equation.] By assuming (as is usual) the orthogonality of $w_{lk}(\mathbf{r})$ for different l

$$\int w_{l'k}^*(\mathbf{r})w_{lk}(\mathbf{r})dV = \delta_{ll'}, \quad (34)$$

where the integration is over the volume of the crystal, we get

$$\int\psi_{q'}^{*(l'k')}(\mathbf{r})\psi_q^{(l,k)}(\mathbf{r})dV = \delta_{qq'}\delta_{kk'}\delta_{ll'}. \quad (35)$$

Hence, the functions (33) form a complete orthonormal set.

IV. INFINITE MAGNETIC TRANSLATION GROUP

The construction of the representations of (23) of $\bar{G}$ and of the symmetry adapted functions (33) was made possible by imposing the Born-von Karman boundary conditions. We know, however, that in the case of the usual translation group it is possible to write down the representations without using the boundary conditions, in which case the wave vector $\mathbf{k}$ describing the representations varies continuously in the first Brillouin zone. One may expect the same to be valid also in the case of the magnetic translation group. The crucial point in deriving the irreducible representations of $\bar{G}$ was the existence of an invariant Abelian subgroup $\bar{F}$ of the group $\bar{G}$. The infinite magnetic translation group G always has an invariant commutative subgroup F [Eq. (23), Ref. 1] when the magnetic field $\mathbf{H}$ is in a direction of a lattice vector. The restriction imposed on the direction of the magnetic field only is a much weaker one and it leaves the group $\bar{G}$ infinite. Since the Born-von Karman boundary conditions do not appear explicitly in the functions (33), it is easy to obtain a generalization of the functions in the case of an infinite M.T.G. (with the restriction that the magnetic field is in a direction of a lattice vector, say, again in the direction of $\mathbf{a}_3$). We have, however, to distinguish the case

$$\mathbf{h}\neq(4\pi/V)(n/N)\mathbf{a}_3 \quad (36)$$

from the case

$$\mathbf{h}=(4\pi/V)(n/N)\mathbf{a}_3, \quad (37)$$

where n and N have no common factor. The integers n and N are no longer connected with the boundary conditions. In the case when (36) holds the symmetry adapted functions will be

$$\psi_q^{(l_1m_3)}(\mathbf{r}) = \exp\{im_3\mathbf{K}_3\cdot\mathbf{r}\}\exp\{i[q+(1/4\pi)\mathbf{K}_2\cdot\mathbf{r}] \times(\mathbf{a}_2\times\mathbf{h}\cdot\mathbf{r})\}w_{lm_3}(\mathbf{r}+q\mathbf{a}_2), \quad (38)$$

where $0\leq m_3<1$ (continuously) and q takes on integer values from 0 to infinity. For the case (37) we get as in (33):

$$\psi_q^{(l,\mathbf{k})}(\mathbf{r}) = \exp\{i\mathbf{k}\cdot\mathbf{r}\}\exp\{i[q+(1/4\pi)\mathbf{K}_2\cdot\mathbf{r}] \times(\mathbf{a}_2\times\mathbf{h})\cdot\mathbf{r}\}w_{lk}(\mathbf{r}+q\mathbf{a}_2), \quad (39)$$

where

$$q=0,\ 1,\ \cdots,\ N-1\ [(N/2)-1] \text{ for odd (even) } N,$$

$$\mathbf{k}=\frac{m_1}{N}\mathbf{K}_1+\frac{m_2}{N}\mathbf{K}_2+m_3\mathbf{K}_3, \quad (40)$$

$$0\leq m_1,\ m_2<1\ (2) \text{ for odd (even) } N,\quad 0\leq m_3<1,$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and

$$w_{lk}(\mathbf{r}+N\mathbf{a}_2)=\exp\{-iN\mathbf{a}_2\times\mathbf{h}\cdot\mathbf{r}\}w_{lk}(\mathbf{r}) \quad \text{for odd } N, \quad (41)$$

$$w_{lk}[\mathbf{r}+(N/2)\mathbf{a}_2]=\exp\{-i(N/2)\mathbf{a}_2\times\mathbf{h}\cdot\mathbf{r}\}w_{lk}(\mathbf{r}) \quad \text{for even } N.$$

One should note that the appearance of $N/2$ for even N is in agreement with the distinction between even and odd N in the construction of the irreducible representations (23).

There are two interesting points connected with the generalization for an infinite M.T.G. The first point is that by changing the magnetic field *continuously* we can pass from case (37) to case (36). This means that by changing the magnetic field by an amount which can be as small as we like we change the degeneracy from a finite one (given by the number N) to an infinite one. We thus see that the representations are very sensitive to the magnitude of the magnetic field. The relation (37) can also be written in a different way by multiplying both sides by $\mathbf{a}_1\times\mathbf{a}_2$. We get (by using the definition of $\mathbf{h}$):

$$\frac{\mathbf{H}\cdot\frac{1}{2}(\mathbf{a}_1\times\mathbf{a}_2)}{hc/e}=\frac{n}{N}, \quad (42)$$

where h is Planck's constant. In the numerator on the left of (42) we have the flux of the magnetic field through an elementary area enclosed by the vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $-(\mathbf{a}_1+\mathbf{a}_2)$. The quantity hc/e is the elementary quantum of magnetic flux (fluxon) which is of great importance in superconductivity[9] and may have an influence on the significance of potentials in quantum mechanics.[10] From Eq. (42) it follows that the representations of the magnetic group are finite when the ratio of magnetic flux through areas enclosed by any vectors of the Bravais lattice to the elementary fluxon are given by rational numbers. Through Eq. (42) we come also to the second point, namely, when the ratio given by this equation is an integer, all the constant factors in the definition of the elements of the M.T.G. [Eq. (1)] will be unity. In this case the magnetic translation group will be isomorphic to the usual translation group and all the irreducible representations of the first will become one dimensional, and one has an energy band picture as in the Bloch case. In this connection, it is interesting to refer to the work of Azbel[11] who shows that for a special Hamiltonian all physical properties of electrons in a magnetic field are periodic in the magnetic field with a period

$$\mathbf{h}=(4\pi/V)\mathbf{a}_3. \quad (43)$$

The relation (43) expressed in the form (42) is just a requirement on the magnetic flux through areas enclosed by any Bravais lattice vectors to be multiples of the fluxon hc/e. The fields satisfying condition (43) must be of the order of 10^{10} G for $a\approx10^{-8}$ cm, and cannot be practically achieved. The fundamental significance of flux quantization in the problem under discussion is, however, of great interest and requires further investigation.

V. CONCLUSION

We have indicated the usefulness of the magnetic translation group for investigating the dynamics of a Bloch electron in a magnetic field. From the irreducible representations of the M.T.G. a set of orthonormal functions was constructed. It turns out that when a magnetic field is present the imposition of boundary conditions leads to quantization of the field, i.e., by imposition of boundary conditions on crystals, even of dimensions of the order of 1 cm, the magnetic field can only take values which differ by about 10 G. (It has been usual to assume that the Born-von Karman boundary conditions have no influence on the physical description of an electron in a periodic potential.[12]) It was therefore important for us to remove the boundary conditions. By doing so, we found the irreducible representation of the M.T.G. (and the symmetry adapted functions) when the magnetic field is required to be along a lattice vector—without any restrictions on the magnitude of the magnetic field. It turns out that the representations are very sensitive to whether or not condition (42) holds.

It would be interesting to find the representations of the M.T.G. without any requirements on the magnetic field at all.

ACKNOWLEDGMENTS

The author would like to thank the National Magnet Laboratory for its very kind hospitality and especially Dr. Arthur J. Freeman for helpful discussions and valuable suggestions in preparing the manuscript. The author would also like to thank Professor G. F. Koster for reading the manuscript and for his encouragement.

[9] M. Peshkin and W. Tobocman, Phys. Rev. 127, 1865 (1962) and references given therein.

[10] Y. Aharonov and D. Bohm, Phys. Rev. 115, 485 (1959); 123, 1511 (1961).

[11] M. Y. Azbel, Zh. Eksperim. i Teor. Fiz. 44, 980 (1963) [English transl.: Soviet Phys.—JETP 17, 665 (1963)].

[12] It is interesting to note that no boundary conditions can be introduced for different N_1, N_2, N_3 (in directions $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ respectively) that have no common factor. In this case no ray representations exist for the usual translation group.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Nuclear Magnetic Relaxation of the Impurity Nucleus in Dilute Ferromagnetic Alloys

R. L. STREEVER
National Bureau of Standards, Washington, D. C.
(Received 27 January 1964)

We have measured the nuclear relaxation times T_1 and T_2 for 1% Ni^{61} in 99% iron, 1% Ni^{61} in 99% cobalt, and 0.5% Co^{59} in 99.5% nickel at several temperatures using the pulsed-free precession method. The relaxation curves are generally nonexponential and power-dependent. The signals at low-power levels are from nuclei in domain walls and T_1 is due to thermal fluctuations of the domain walls. At high-power levels, where the signal is mainly from nuclei in domains, the longest measured relaxation times T_1 are lower limits for relaxation times for nuclei in domains. The longest T_1's are found to be inversely proportional to temperature with $T_1T=1.0$ sec °K for 1% Ni^{61} in cobalt, $T_1T=1.2$ sec °K for 1% Ni^{61} in iron, and $T_1T=0.3$ sec °K for 0.5% Co^{59} in nickel, and are believed to be due to conduction electron relaxation. The quantity $1/\gamma_n^2T_1T$ (where γ_n is the nuclear gyromagnetic ratio) is a normalized measure of the strength of the conduction electron relaxation mechanism and is found to be smaller for the impurity nuclei than for the nuclei of the solvent atoms in the pure solvent metal. This is believed to show that the relaxation of the nuclei by the conduction electrons (which is enhanced in the pure metals by spin waves) is slower for the nuclei on the solute atoms in the alloys due to a reduction of the spin wave enhancement. T_2 for Ni^{61} in cobalt is 0.15 msec at room temperature and is believed to be due to a coupling of Ni^{61} and Co^{59} nuclei by the Ruderman-Kittel interaction.

I. INTRODUCTION

In the pure metals, iron, cobalt, and nickel, the nuclear relaxation times T_1 and T_2 have been measured by Weger[1–3] by the free precession method. The decay of the longitudinal and transverse magnetization in the ferromagnetic metals is found to be nonexponential and a function of the rf power level. Nuclei in the domain walls experience a large enhancement of the applied rf level due to the domain wall motion. Thermal fluctuations of the domain walls also provide a relaxation mechanism; consequently, at low rf pulse levels, the nuclear signals are from nuclei near the center of the domain walls which have the largest rf enhancement and shortest relaxation times. At higher rf pulse amplitudes signals come from nuclei farther out from the centers of domain walls, which have lower enhancements of rf fields and longer relaxation times. Finally, at very high rf pulse levels, signals come primarily from nuclei in the domains which see a much smaller rf enhancement than those in the domain walls. For these nuclei, relaxation times T_1 are longer, being determined by conduction electrons now and no longer by domain wall motions. For T_2, the experimental situation is more complicated. For low abundance of nuclear spins and at high temperatures the nuclear spin-spin interactions can generally be neglected. In this case, T_2 is found to be of the same order as T_1. For cases in which nuclear spin-spin interactions are important, as in pure cobalt, these determine T_2.

In the present work we have used the free precession method to measure T_1 and T_2 for Ni^{61} in dilute alloys of 1% Ni^{61} in cobalt[4] and iron,[5] and for Co^{59} in a 0.5% Co^{59} in nickel alloy.[6] Preliminary to studying relaxation times of the impurity nuclei in the alloys, some measurements were also made of T_1 and T_2 for Ni^{61} in pure unenriched nickel powders to determine the effect of annealing on the relaxation times. The Ni^{61} relaxation times were also measured in the 0.5% cobalt in nickel alloy to determine the effect of the cobalt impurity on the Ni^{61} relaxation.

In Sec. II the experimental methods and techniques are discussed. In Sec. III we review briefly the theory of relaxation, both in domain walls and in the domains. In Sec. IV we give the experimental results, and compare them with the theory.

II. EXPERIMENTAL METHOD AND TECHNIQUE

A. Equipment

We have used the free precession method to measure the inhomogeneous linewidth and the longitudinal and transverse relaxation times T_1 and T_2. A pulsed push-pull rf oscillator (Arenberg Ultrasonic Laboratories Model 650-C) was used to supply the rf pulses. The widths of the rf pulses used were about 1 μsec and the maximum rf voltage level was about 300 V peak-to-peak rf. The metal powder was inserted in the rf coil of the oscillator. A second coil coaxial to the first was used to detect the free precession signal. It was convenient to heterodyne the nuclear signal with a mixing frequency differing by about 20 Mc/sec, using a television tuner. The television tuner was followed by a wide-band rf

[1] M. Weger, thesis, University of California, Berkeley, 1961 (unpublished).
[2] M. Weger, E. L. Hahn, and A. M. Portis, J. Appl. Phys. 32, 1645 (1961).
[3] M. Weger, Phys. Rev. 128, 1505 (1962).
[4] R. L. Streever, L. H. Bennett, R. C. La Force, and G. F. Day, Phys. Rev. 128, 1632 (1962).
[5] R. L. Streever, L. H. Bennett, R. C. La Force, and G. F. Day, J. Appl. Phys. 34, 1050S (1963).
[6] L. H. Bennett and R. L. Streever, J. Appl. Phys. 33, 1093S (1962).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

amplifier and detector. The video signal was then amplified and observed directly on the oscilloscope. The dc pulses necessary to gate the pulsed oscillator were supplied by appropriate combinations of Tektronix Nos. 162 and 163 units.

B. Measuring Methods

To measure T_1 and T_2, one usually employs a method in which one makes use of 90° and 180° pulses. In a ferromagnetic metal as a result of the skin depth and the distribution in enhancement factors, different nuclei see different rf fields. Consequently, 90° and 180° pulses are impossible to obtain. The decay of the magnetization in the transverse direction with time constant T_2 is, however, independent of pulse adjustments and T_2 can be obtained from the echo decay envelope. To measure T_1, we saturated the nuclear signal with one or more closely spaced rf pulses, and then monitored the recovery of the magnetization by observation of the echo following two closely spaced pulses applied some time after the saturating pulses. An echo rather than a single pulse was used to sample the recovery of the magnetization because for the broad resonance lines where the echo width is a few μsec, the receiver blocking after the rf pulse makes the decay itself impossible to observe. Weger[1] has shown that the above method is subject to some error due to diffusion of magnetization in space out of the domain walls. By applying a large number of saturating pulses whose separation is much less than T_2, this effect can to a large extent be overcome.

The relaxation times T_1 and T_2 were measured as a function of rf pulse amplitude. The echo signal following two pulses of equal width goes from zero through a maximum at low rf pulse amplitudes as the pulse amplitude is increased from zero, then slowly decreases again as the pulse amplitude is further increased, as discussed in Ref. 1. The variation in the echo signal intensity results from the distribution in rf fields due to skin depth and a distribution in enhancement factors, and the signal reaches a maximum when $\gamma_n\eta_1 H_1 t_w \sim \pi/2$. Here η_1 is an *average effective* enhancement factor, $2H_1$ is the linearly applied rf field, t_w is the pulse width, and γ_n is the nuclear gyromagnetic ratio. The relaxation curves were in general nonexponential, and a function of the amplitude of the rf pulses for a given pulse width, which was usually about one μsec. T_2 and T_1 can in the case of a nonexponential decay be defined for any point on the relaxation curves from the slope. That is,

$$1/T_2 = -d\ln M(t)/dt,$$

where $M(t)$ is the amplitude of the echo signal at a time t after the first pulse and

$$1/T_1 = -d\ln[M(\infty)-M(t)]/dt,$$

where $M(t)$ is the value of the sampling signal at a time t after the last saturating pulse.

C. Samples

The majority of the relaxation time studies were made on alloy powders of 1% Ni^{61} and 99% cobalt,[4] 1% Ni^{61} and 99% iron,[5] and 0.5% Co^{59} and 99.5% nickel[6] (all atomic percents). The Ni^{61} in cobalt was annealed for $\frac{1}{2}$ h at 600°C. The Ni^{61} in iron sample was annealed for $\frac{1}{2}$ h at 680°C. The Co^{59} in nickel sample was annealed for 1 h at 1100°C, but some measurements were also made on an unannealed sample. The relaxation times for nuclei in the domain walls is expected to depend somewhat on the degree of particle annealing and will be discussed in Sec. III.A. The average particle size was about 10 μ. More details of sample preparation have been previously discussed.[4–6]

Some studies were also made on the Ni^{61} resonance in pure unenriched nickel samples with various degrees of annealing to investigate the effects of strains on the relaxation times. These samples were all 10 μ powders. The preparation of the pure nickel samples has been discussed previously.[7]

III. THEORY

A. Relaxation in the Domain Walls

Nuclear relaxation in the domain walls has been studied both theoretically and experimentally by several authors.[1,3,8–10] For the pure metals Weger[3] gives for $1/T_1$

$$1/T_1 = (kT\tau_c\omega^2/\delta^2\pi M_s^2)(1+\omega^2\tau_c^2)^{-1}. \quad (1)$$

Here k=Boltzmann constant, T=absolute temperature, τ_c=correlation time for wall motion due to thermal fluctuations, δ=wall thickness, M_s=saturation magnetization, and ω=resonance frequency. For cobalt at room temperature Eq. (1) gives good agreement with experiment.[3] For the pure metals and the alloys studied in this paper, $\omega\tau_c<1$ is a good approximation. In this limit

$$1/T_1 = kT\tau_c\omega^2/\pi\delta^2 M_s^2. \quad (2)$$

The weak temperature dependence of T_1 in the pure metals is believed to arise because while kT enters directly into Eq. (2), τ_c probably increases with decreasing temperature due to greater eddy current damping of the domain wall motion.[3]

Equation (2) was derived for the pure metals; however, quite generally in the limit $\omega\tau<1$ the relaxation time will be given by (see Ref. 11)

$$1/T_1 = \langle H_f^2\rangle\gamma_n^2\tau_c. \quad (3)$$

Here H_f is the fluctuating field perpendicular to the hyperfine field H_n resulting from thermal fluctuations

[7] R. L. Streever and L. H. Bennett, Phys. Rev. **131**, 2000 (1963).
[8] J. M. Winter, Phys. Rev. **124**, 452 (1961).
[9] E. Simanek and Z. Sroubek, Czech. J. Phys. **311**, 764 (1961).
[10] M. Matsuura *et al.*, J. Phys. Soc. Japan **17**, 1147 (1962).
[11] C. P. Slichter, *Principles of Magnetic Resonance* (Harper and Row, New York, 1963), Chap. 5, p. 153.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

of the domain wall. Equation (3) can be rewritten

$$1/T_1=\langle\theta^2\rangle\gamma_n^2H_n^2\tau_c=\langle\theta^2\rangle\tau_c\omega^2, \quad (4)$$

where θ is the angular displacement of the electron spin in the domain wall.

At a given temperature Eq. (4) shows that either an increase in the amplitude of the wall motion or the correlation time of the motion will shorten T_1. These parameters should be sensitive to the degree of anneal of the metal. Hence, in the pure metals, for example, in pure nickel, we might expect the Ni^{61} relaxation times T_1 to depend on the degree of annealing of the sample. Also, for the same annealing time we might expect the parameters $\langle\theta^2\rangle$ and τ_c to depend on the concentration of impurities so that the relaxation times T_1 of the solvent nuclei might be expected to depend on the impurity concentration. We also see from Eq. (4) that samples with the highest rf enhancements (due to domain wall motion) will be expected to have the shortest relaxation times for the nuclei in domain walls, a result previously noted.[10]

Consider now the relaxation time of the nuclear spin of the solute atom. From Eq. (4) we see that if the impurity spin in the domain wall moves with the same correlation time and angular displacement as the solvent atom spin in the dilute alloy, then the value of $T_1\omega^2$ for the nuclei of the solute and solvent atoms should be equal. This relation is used in Sec. IV.C to discuss values of T_1 measured for the nuclei of the solute atoms in the domain walls.

B. Relaxation in the Domains

The spin-lattice relaxation time for nuclei in the domains has been considered by Weger.[3] The relaxation time T_1 for nuclei in domains due to conduction electron relaxation is given by the expression

$$1/T_1=(kT\omega^2a^2\Sigma)(h\omega_d^2 32\pi^3S^2)^{-1}, \quad (5)$$

where a=lattice constant, ω_s=parameter describing spin-wave spectra, S=average spin/atom, Σ=area of Fermi surface, ω=nuclear resonance frequency, k=Boltzmann constant, and T=absolute temperature.

Equation (5) applies to the pure metals and the T_1 mechanism is a second-order process which involves the coupling of the nuclei to the unpaired electron spins through the hyperfine interaction ($A\mathbf{I}\cdot\mathbf{S}$) and the coupling of the electron spins to the conduction electrons through the *sd* coupling. The intermediate states involve the excitation of spin waves.

In the cases where hyperfine fields are proportional to local atomic moments (this is, at least, roughly true for the cases of concern here),[5] the factor (ω^2/S^2) appearing in Eq. (5) is just proportional to γ_n^2, where γ_n is the nuclear gyromagnetic ratio. Quite generally, the quantity $1/\gamma_n^2T_1T$ is a normalized measure of the strength of the conduction electron relaxation mechanism. If the mechanism in which the nuclear relaxation is enhanced by the spin waves applies to the solute nuclei in the dilute alloys, as well as to the pure metals, then the quantity $\gamma_n^2T_1T$ would be expected to be of the same order of magnitude for the nucleus of the solute atom in the dilute alloy as for the nuclei of the solvent atoms in the same alloy or in the pure solvent metal. One might expect that the creation of spin waves by a spin flip of the solute atom will in general involve a higher energy than in the case of the same process in the pure metal and that the enhanced conduction electron relaxation mechanism will be weaker for the solute nuclei. In this case the quantity $\gamma_n^2T_1T$ will be longer for the solute nuclei and closer to its value in the nonferromagnetic materials.

C. Transverse Relaxation

Weger has measured T_2 as well as T_1 in the pure metals; iron, cobalt, and nickel.[1–3] In the case of Ni^{61} or Fe^{57} in the pure unenriched metals, where nuclear spin-spin interactions are weak, T_2, at least at higher temperatures and in the domain walls, is of the order of T_1 and the transverse relaxation is related to the T_1 mechanism. In the case of Co^{59} in pure cobalt where the abundance of active nuclei is high, T_2 is determined by nuclear spin-spin interactions.

At low temperatures in the domains where T_1 is long, T_2 may be determined by a mechanism of the type discussed by Moriya.[12] He considers in detail nuclear relaxation in antiferromagnets; however, similar considerations would apply to ferromagnets as well. Fluctuations in S_z can arise from the interference of two spin waves and might be particularly important in the alloys. These fluctuations would be very effective for relaxation since they occur at lower frequencies than the spin waves. Hence, at low temperatures in the domains, we might expect $T_2<T_1$, even in cases where the nuclear spin-spin coupling can be neglected.

IV. RESULTS AND DISCUSSION

A. T_1 and T_2 for Ni^{61} in Pure Nickel and in the Dilute Alloys

Preliminary to studying relaxation times of the solute nuclei in the dilute alloys, measurements were made of T_1 and T_2 for Ni^{61} in pure unenriched nickel powder with a particle size of 10 μ or less to determine the effect of annealing. The measurements were made at 77°K and the initial relaxation rates were measured at low pulse levels to determine T_1 and T_2 appropriate to the nuclei in domain walls. We found for unannealed samples or for samples annealed for 1 h at 300 and 500°C that T_1 was roughly independent of annealing and about 1.2 msec with T_2 about 0.3 msec. In nickel annealed at 1100°C for 1 h, however, T_1 decreased to 0.5 msec and T_2 to 0.06 msec. The Ni^{61} relaxation time

[12] T. Moriya, Progr. Theoret. Phys. (Kyoto) 16, 641 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

appropriate to nuclei in domain walls was also measured at 77°K in a sample of 99.5% nickel and 0.5% cobalt, which was annealed at 1100°C for 1 h. T_1 was found to be 1.3 msec, which is about the same as the T_1 in the pure nickel which was not well annealed.

The decrease in T_1 in the well-annealed samples can be explained as follows: In the well-annealed nickel, the initial permeability is high and we expect a larger amplitude for the domain wall motion (for a given excitation). This is in agreement with previous continuous-wave studies,[7] where the large dispersion signals in the well-annealed samples were explained partly by the larger rf enhancements in these samples. The decrease in T_2 is partly related to the T_1 decrease; however, the ratio of T_1 to T_2 is even greater in the well-annealed sample, indicating that fluctuations in S_z along the direction of the nuclear hyperfine field may be more important here. We also see that in the well-annealed sample of 0.5% cobalt in nickel the Ni^{61} relaxation time T_1 is about the same as in the unannealed pure nickel. Both strains as well as impurities seem to have the effect of limiting the displacement of the domain walls and lengthening the relaxation time. For the same annealing time at least the relaxation time is longer in the 0.5% cobalt sample than in the pure nickel.

B. T_1 and T_2 of the Solute Nuclei in the Alloys

The longitudinal and transverse relaxation curves were obtained as a function of pulse level at various temperatures, as discussed in Sec. II.B, for the solute nuclei in annealed samples of 1% Ni^{61} in cobalt, 1% Ni^{61} in iron and 0.5% Co^{59} in nickel. The relaxation curves were, in general, nonexponential and dependent on the pulse level. The longest and shortest times are given in Table I. The shortest times T_{10} and T_{20} were determined from the initial slopes of the respective relaxation curves obtained at a low pulse amplitude below that at which the echo signal went through a maximum, and are believed to approach the shortest relaxation times appropriate to nuclei near the center of the domain walls. The longest times T_{1M} and T_{2M} were measured at the highest pulse level from the slope of the respective relaxation curves at the farthest time out, and are believed to approach the values for nuclei in the domains. The highest rf pulse level which was roughly 12 times that at which the echo signal went through a maximum corresponded to an H_1 of roughly 6 G for Ni^{61} in cobalt and Ni^{61} in iron and about 3 G for Co^{59} in nickel.

TABLE I. Relaxation times for the impurity nuclei in the dilute alloys, T_{10} and T_{20} are the relaxation times appropriate to nuclei in domain walls, while T_{1M} and T_{2M} approach the relaxation times in domains as discussed in the text.

Temp. (°K)	T_{10} (msec)	T_{1M} (msec)	T_{20} (msec)	T_{2M} (msec)
		1% Ni^{61} in Cobalt		
300	1.4	3.5	0.15	0.15
77	2.0	13	0.18	0.18
4	3.5	250	0.19	0.22
		1% Ni^{61} in Iron		
300	0.7	4	0.17	2.0
77	1.7	15	0.6	2.4
4	6.0		1.0	3.0
		0.5% Co^{59} in Nickel		
300	0.05	1	0.03	0.12
77	0.13	4	0.04	0.18
4	1.4		0.15	0.34

The longitudinal relaxation curves were dependent on the rf pulse level, nonexponential at low pulse levels and more nearly exponential at higher pulse levels, where the signal was mainly from nuclei in domains. For the case of Ni^{61} in iron and Co^{59} in nickel, the transverse decay curves at the higher temperatures were also dependent on the pulse level and nonexponential and T_2 in the domain walls was of the same order as T_1, indicating it was related to the T_1 mechanism. At lower temperatures, particularly in the domains, T_2 is seen from Table I to be much less than T_1 and less dependent on pulse level. In the case of Ni^{61} in cobalt, the transverse decay curves are nearly exponential and T_2 was nearly independent of pulse level. The nearly unique value of T_2 for this case indicates it is being determined by nuclear spin-spin interactions, a similar result having been observed for the case of Co^{59} in pure cobalt metal.[1]

The inhomogeneous linewidths at room temperature obtained from the width of the echo (assuming an exponential decay) and measured by the quantity $1/\pi T_2^*$ were about 240 kc/sec for 1% Ni^{61} in cobalt and 190 kc/sec for 0.5% Co^{59} in nickel and were in approximate agreement with continuous wave measurements. The 1% Ni^{61} in iron width (~700 kc/sec) was too broad to estimate accurately from the width of the echo.

C. Discussion of T_1 in Domain Walls

In Table II we have compared the value of T_{10} (the T_1 appropriate to nuclei in domain walls) for the nuclei in the dilute alloys with the shortest T_1's measured in the pure solvent metals by Weger[2] which correspond to our T_{10}'s. In view of Eq. (4), if the angular displacement and the correlation time associated with the thermal fluctuations of the solute electron spin in the domain wall were the same as for the electron spin in the pure solvent metal, then the quantity $T_{10}\nu^2$ (where ν is the resonance frequency) should be the same for the solute nucleus as for the nucleus in the pure solvent metal. In the last two columns of Table II we compare the quantity $\nu^2 T_{10}$ for the solute nuclei in the alloys and the nuclei in the pure solvent metal. We see that in all cases this quantity is longer for the solute nuclei. This can probably be understood as due to the fact that the thermal fluctuations of the domain wall in the dilute

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. A comparison of the longitudinal relaxation times in the domain walls, T_{10}, for the impurity nuclei in the dilute alloys (from Table I) with the corresponding times in the pure host metal measured by Weger *et al.*[a] The quantities ν^2T_{10} are also compared.[b]

Temp (°K)	T_{10} (msec)[a] Co^{59}	T_{10} (msec) Ni^{61}	ν (Mc/sec) Co^{59}	ν (Mc/sec) Ni^{61}	ν^2T_{10} (Mc²/sec) Co^{59}	ν^2T_{10} (Mc/sec) Ni^{61}
	T_{10} for Co^{59} in pure cobalt and 1% Ni^{61} in 99% cobalt.					
300	0.1	1.4	213	70	4.5	6.9
77	0.2	2.0	217	72	9.4	10.4
4	0.2	3.5	217	72	9.4	18.1
	T_{10} for Fe^{57} in pure iron and 1% Ni^{61} in 99% iron.					
Temp (°K)	T_{10} (msec)[a] Fe^{57}	T_{10} (msec) Ni^{61}	ν (Mc/sec) Fe^{57}	ν (Mc/sec) Ni^{61}	ν^2T_{10} (Mc²/sec) Fe^{57}	ν^2T_{10} (Mc/sec) Ni^{61}
300	0.9	0.7	45	85	1.8	5.1
77	0.7	1.7	47	89	1.5	13.5
4	10.0	6.0	47	89	22.1	47.5
	T_{10} for Ni^{61} in pure nickel and 0.5% Co^{59} in 99.5% nickel.					
Temp (°K)	T_{10} (msec)[a] Ni^{61}	T_{10} (msec) Co^{59}	ν (Mc/sec) Ni^{61}	ν (Mc/sec) Co^{59}	ν^2T_{10} (Mc²/sec) Ni^{61}	ν^2T_{10} (Mc²/sec) Co^{59}
300	0.35	0.05	26	112	0.24	0.63
77	1.5	0.13	28	120	1.2	1.9
4	15	1.4	28	121	11.8	20.5

[a] See Ref. 2.
[b] The resonance frequencies are rounded to the nearest megacycles. See A. C. Gossard, Ph.D. thesis, University of California, 1960 (unpublished); J. I. Budnick, L. J. Bruner, R. Blume, and E. L. Boyd, J. Appl. Phys. 32, 120S (1961) and Refs. 4–7.

alloy tend to be decreased compared with those in the pure solvent metal.

Although in Table II the T_{10} for the solute nuclei is compared with the T_{10} in the pure host nickel, it would be better to actually compare the relaxation times for the nuclei of the solute and solvent atoms in the same sample. In the case of 0.5% cobalt in nickel at 77°K we have done this. As already discussed (see Sec. IV.A), T_{10} for the Ni^{61} in a sample annealed for one hour at 1100°C was 1.3 msec, while for Co^{59} the value of T_{10} in the same sample was 0.13 msec. From Eq. (4) we see then that $\langle\theta^2\rangle\tau_c$ is about one-half as large for the Co^{59} as for the Ni^{61}, and this probably represents a somewhat smaller angular displacement of the cobalt spin in the domain wall.

In the 0.5% cobalt in nickel sample, the Co^{59} relaxation times in domain walls were found to be nearly independent of annealing time. This seems to indicate that in this sample the solute atoms themselves rather than the degree of anneal are limiting the wall motion.

D. Discussion of T_1 in the Domains

The longer times T_{1M} given in Table I are nearly inversely proportional to temperature and are believed to approach the limiting value of T_1 appropriate to the nuclei in domains and to be determined by relaxation through the conduction electrons. The quantity $1/\gamma_n^2T_{1M}T$ is a measure of the strength of the conduction electron relaxation mechanism normalized for the difference in nuclear gyromagnetic ratios. In Table III we compare $\gamma_n^2T_{1M}T$ for the nuclei on the impurity atoms with the corresponding quantity for the nuclei of the pure host metal using the value of T_1T measured in the pure metals by Weger.[3] We also list in the same table the atomic magnetic moments measured by neutron diffraction.[13,14]

The quantity $\gamma_n^2T_{1M}T$ is seen in all cases to be greater for the impurity nucleus than for the nucleus of the corresponding pure host metal and this is particularly true for the case of 0.5% Co^{59} in nickel. Also, especially in the case of Ni^{61} in iron where the signal at high power levels was weak and the dependence of T_1 on the power level was quite strong, we cannot be sure we have measured the longest time so that the T_{1M} is only a lower limit for the value of T_1 in the domains.

The longer relaxation times for the impurity nuclei can probably be understood qualitatively as representing a partial decoupling of the spin on the impurity

TABLE III. A comparison of the quantity $T_{1M}T$ for the impurity nuclei in the dilute alloys (from Table I) with the corresponding quantity in the pure host metal, as measured by Weger.[a] The quantities $\gamma_n^2T_{1M}T$ are also compared.[b] The atomic moments[c] in Bohr magnetons (bm) are also given.

Nucleus	μ (bm)	$T_{1M}T$ (sec°K)	γ_n (cps G^{-1})	$\gamma_n^2T_{1M}T$
Co^{59} in Co	1.7	0.08	6.27×10^3	3.1×10^6
1% Ni^{61} in 99% Co	0.6	1.0	2.38×10^3	5.7×10^6
Fe^{57} in Fe	2.2	2.5	0.864×10^3	1.9×10^6
1% Ni^{61} in 99% Fe	1.0	1.2	2.38×10^3	6.8×10^6
Ni^{61} in Ni	0.6	0.115	2.38×10^3	0.65×10^6
0.5% Co^{59} in 99.5% Ni	1.7	0.3	6.27×10^3	11.8×10^6

[a] See Ref. 3.
[b] For values of γ_n see Freeman *et al.*, Proc. Roy. Soc. (London) 242, 455 (1957) for Co^{59}; P. R. Locher and S. Geschwind, Phys. Rev. Letters 11, 333 (1963) for Ni^{61}; and C. W. Ludwig and H. H. Woodbury, Phys. Rev. 117, 1286 (1960) for Fe^{57}.
[c] See Refs. 13 and 14.

[13] M. F. Collins and J. B. Forsyth, Phil. Mag. 8, 401 (1963).
[14] M. F. Collins and D. A. Wheeler, Proc. Phys. Soc. (London) 82, 633 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

atom from the spin-wave spectrum; that is, a spin flip of the impurity atom spin, which is different from that of the host (see Table III), will probably in general require a higher energy than for the case of the pure metal. This probably reduces the enhancement of the conduction electron relaxation by the spin waves.

Note also that while $\gamma^2T_{1M}T$ for Ni^{61} in cobalt and iron is longer than in pure nickel, it is still shorter than γ^2T_1T for the case of a nonferromagnetic metal like Cu^{63} in copper where γ^2T_1T is about 45×10^6 (with γ in cpsG^{-1} and T_1T in sec °K).[15] Similarly, while $\gamma^2T_{1M}T$ for Co^{59} in nickel is longer than for Co^{59} in pure cobalt, it is still less than γ^2T_1T for V^{51} in vanadium, where γ^2T_1T is about 39×10^6.[16]

E. Discussion of T_2

In the case of Ni^{61} in cobalt, T_2 is being determined by nuclear spin-spin interactions. We will discuss various coupling mechanisms. The Suhl-Nakamura interaction[17,18] involving virtual excitation of spin waves is a nuclear spin-spin coupling mechanism for Co^{59} in pure cobalt.[1] This mechanism, however, at least in the pure ferromagnets and at low temperatures (well below the Curie point), couples nuclear spins according to $I_1^+I_2^-+I_2^+I_1^-$ and this would not be a broadening mechanism for unlike nuclear spins which would not be precessing at the same Larmour frequency.[18] This would probably to a large extent be true also for the case of 1% Ni^{61} in cobalt, where the spin-wave excitations are still mainly transverse.

The Ruderman-Kittel[19] interaction involving a scalar coupling between nuclear spins would be a broadening mechanism for unlike spins. One might argue that the same mechanism which leads to the short values of T_1T for both Co^{59} and Ni^{61} in the 1% Ni^{61} in cobalt alloy would also lead to an enhanced Ruderman-Kittel interaction. However, an enhanced interaction of this sort involving spin-wave excitations would at low temperatures be anisotropic in the same way as the Suhl interaction and would not be effective in broadening unlike spins. However, the ordinary Ruderman-Kittel interaction is about the right value to explain the observed T_2. This can be seen as follows: In the case of Cu^{63} in copper metal, the experimental results could be explained by taking the scalar indirect exchange coupling A/h between two nearest Cu^{63} neighbors to be about 420 cycles or three times the calculated value.[15] Assuming a similar coupling between the nickel and cobalt nuclei leads to an exchange coupling between neighbors of

$$420\gamma(Co^{59})\gamma(Ni^{61})/[\gamma(Cu^{63})]^2=120 \text{ cycles}.$$

The nearest-neighbor contribution to the Van Vleck second moment is then given by

$$\langle\Delta\nu^2\rangle=4I(I+1)(A/h)^2.$$

This gives a width $[\langle\Delta\nu^2\rangle]^{1/2}$ of about 1 kc/sec or T_2 about 0.2 msec. Although the above is the order of magnitude of the T_2 observed, the fast spin exchange of the Co^{59} nuclei would tend to narrow the line and make the shape more nearly Lorentzian so that an exact calculation of T_2 is difficult.

In the case of Ni^{61} in iron and Co^{59} in nickel (where we believe we can neglect nuclear spin-spin interaction) the T_2 relaxation mechanism is evidently related to T_1 at high temperatures in the domain walls. At low temperatures, particularly in the domains, T_2 may be due to the mechanism discussed by Moriya. However, the situation is not completely understood.

V. CONCLUSION

The longitudinal relaxation times for the nuclei at the solute atoms in the domain walls in these alloys can be understood as due to the motion of the domain walls. However, it appears that the relaxation time for the solute nuclei will generally be somewhat longer than the relaxation time for the nucleus at the host atom due to the pinning of the domain wall near the solute atom. In the domains it appears that the coupling of the nucleus to the conduction electron which determines T_1 at high power levels is weaker for the solute atom than for the host atom in these materials. In the case of Ni^{61} in cobalt, T_2 is being determined by nuclear spin-spin interaction. A Ruderman-Kittel indirect coupling of nuclei by conduction electrons seems to account for the observed value of T_2.

ACKNOWLEDGMENTS

The author would like to gratefully acknowledge the experimental assistance of D. E. Brown and G. A. Uriano.

[15] A. G. Redfield, Phys. Rev. **98**, 1787 (1955).
[16] J. Butterworth, Phys. Rev. Letters **5**, 305 (1960).
[17] H. Suhl, J. Phys. Radium **20**, 333 (1959).
[18] T. Nakamura, Progr. Theoret. Phys. (Kyoto) **20**, 542 (1958).
[19] M. A. Ruderman and C. Kittel, Phys. Rev. **96**, 99 (1954).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Further Results on the Fermi Surface of Beryllium*

T. L. Loucks†
North American Aviation, Tulsa, Oklahoma
(Received 3 January 1964)

Following the methods presented in an earlier paper, the conduction eigenvalues for beryllium at the equivalent of over 80 000 points in the first Brillouin zone (all in the immediate vicinity of the Fermi surface) were calculated. The constant energy surfaces were constructed for several values of the energy near the Fermi energy. The hole and electron volumes were calculated for each case, and the Fermi energy was determined by the requirement that the two volumes be equal. This is the first time that this well-known result has actually been used to determine the Fermi energy in a band calculation. The corresponding Fermi surface was defined by about ten cross sections perpendicular to the [0001] direction. The agreement between these results and experiment is generally good and slightly improved over the previous results.

INTRODUCTION

IN a previous article[1] the Fermi surface of beryllium was determined theoretically and compared with the experimental results of Watts.[2] The surface was constructed from conduction electron eigenvalues at the equivalent of 5184 general points in the first Brillouin zone (BZ). These eigenvalues were calculated by expanding the wave functions in a linear combination of 23 orthogonalized plane waves (OPW's). The crystal potential used in these calculations was self-consistent and its construction was fully described in the original article.

The same program which was developed for the earlier calculations has been used to calculate the conduction electron eigenvalues at the equivalent of over 80 000 points in the first BZ. All of these points were in the immediate vicinity of the Fermi surface. From these eigenvalues the Fermi energy was determined for the first time by the well-known requirement that the hole and electron volumes be equal. The Fermi surface corresponding to this Fermi energy was constructed and is presented in this report. It differs only slightly from the previous results, but tends to bring the theoretical and experimental results into closer agreement. The interesting feature here is not the new results, but rather the reminder that there is available, through the application of a well-known conservation requirement, a method of determining the Fermi energy which requires the calculation of energy eigenvalues only in the vicinity of the Fermi surface.

DETERMINING THE FERMI ENERGY

In the work of Loucks and Cutler[1] the Fermi energy was determined in the usual way by arranging in increasing value the energy eigenvalues which were calculated for representative points throughout the BZ. The energy corresponding to the highest occupied state was taken to be the Fermi energy. This method, although correct in principle, suffers in that it is difficult to get a representative sample of points (particularly near the zone edges) without calculating the eigenvalues at an extremely large number of points throughout the zone. Many of these points are not near the Fermi surface and hence are of little value except in the counting procedure for the determination of the Fermi energy.

The method proposed here has the advantage that all eigenvalues are calculated at points in the immediate vicinity of the Fermi surface. Thus, they serve not only to determine the Fermi energy but also to define more accurately the surface itself. In addition it provides enough information about the other energy surfaces near the Fermi surface to enable one to calculate first and second derivatives for use in determining such parameters as the density of states at the Fermi energy and the effective mass.

Essentially, the method employed is based on the obvious conservation requirement that the hole volume in the first double zone must equal the volume of the electron pockets in the second double zone. Since the electrons spill out of one zone into the next in order to

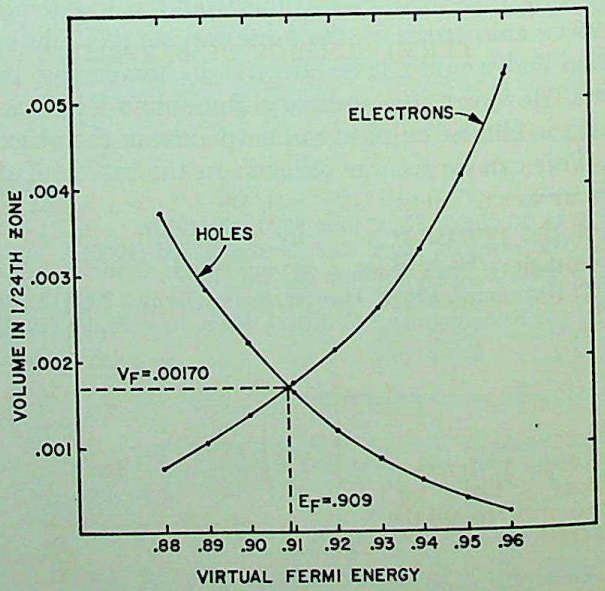


FIG. 1. Dependence of electron and hole volume on virtual Fermi energy.

* Much of this work was done while the author was at the Pennsylvania State University, University Park, Pennsylvania, and supported by the U. S. Air Force Office of Scientific Research, Grant Number AF-AFOSR-61-100.
† Present address: Iowa State University, Ames, Iowa.
[1] T. L. Loucks and P. H. Cutler, Phys. Rev. 133, A819 (1964).
[2] B. R. Watts, Phys. Letters 3, 284 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

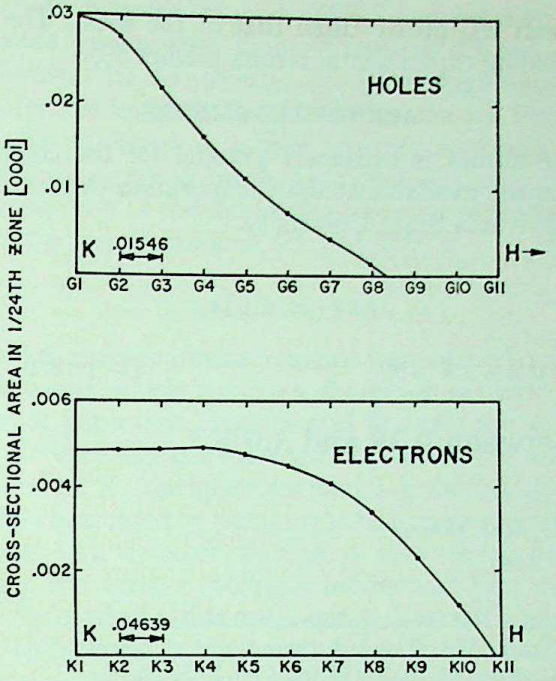


FIG. 2. Cross-sectional area of electron and hole surfaces in 1/24th zone for planes perpendicular to the [0001] direction.

minimize the energy, the vacated volume must be exactly equal to the volume occupied by the electrons. This is an admittedly obvious fact. However, it has never been used in practice to determine a theoretical value of the Fermi energy from a band calculation. The reason it has been avoided is perhaps because of the large number of calculations necessary to determine the electron and hole volumes. In this work they were determined for each of several constant energy surfaces near the Fermi surface. The results are plotted in Fig. 1. The curves of the hole volume and electron volume were found to intersect at $E_F=0.909$ Ry. The Fermi energy presented in the original paper was $E_F=0.901$ Ry.

FERMI SURFACE OF BERYLLIUM

The constant energy surface was constructed at $E_F=0.91$, as well as at other points, in constructing the curves shown in Fig. 1. Since we have found that the true Fermi energy is $E_F=0.909$, it is within the accuracy of the procedure used in constructing the surfaces to consider the one for $E_F=0.91$ to be the Fermi surface. These results are presented in Figs. 2, 3, and 4. In Fig. 2 the cross-sectional area perpendicular to the [0001] direction is shown for the holes and the electrons. In both drawings the left-hand side represents the symmetry point K, and the abscissa corresponds to distances measured along the symmetry edge KH. The identifications such as G1, G2, etc., correspond to the cross sections shown in Fig. 4. It should be noticed that in the lower curve of Fig. 2, corresponding to the electron cigars, the abscissa extends from K to H. However, in the upper curve, corresponding to the hole coronet, the distance between G1 and G11 extends only $\frac{1}{3}$ of the distance from K to H. Figures 3 shows the intersection of the Fermi surface with the 1/24th zone faces.

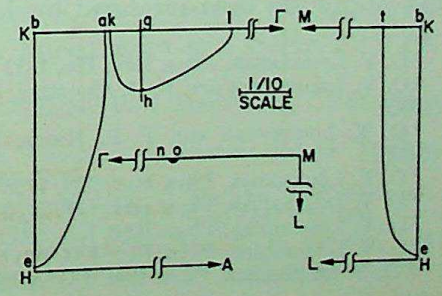


FIG. 3. Intersections of Fermi surface with faces of the 1/24th zone.

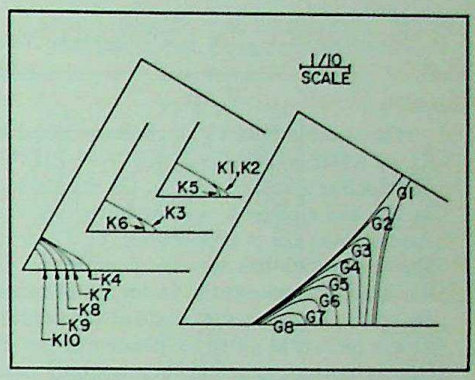


FIG. 4. Intersections of Fermi surface with planes perpendicular to the [0001] direction.

DISCUSSION

The Fermi surface which corresponds to the Fermi energy determined by the above method is compared to the previous theoretical results and the experimental results in Table I. It is seen that the agreement between the present theory and experiment is very good, especially for the bottom four entries which correspond to the coronet. However, the theoretical cigar is still shorter than the experimental one and almost triangular, rather than circular in cross section. This almost triangular shape also yields extremal cross-sectional areas

TABLE I. Comparison of representative dimensions of the Fermi surface as determined by the present work, the de Haas–van Alphen measurements, and the previous theoretical work.

Designation	Loucks	Watts	Loucks & Cutler
be	0.453	0.48	0.44
ba	0.141	0.09	0.13
bt	0.075	0.09	0.07
ak	0.008	0.08	0.01
kl	0.246	0.23	0.26
gh	0.118	0.13	0.13
no	0.017	0.02	0.04
ml	0.574	0.56	0.57

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which are consistent with the observed de Haas–van Alphen frequencies. Further experiments or calculations which can distinguish between this model and the circular one proposed by Watts are necessary. It should be noticed in Fig. 2 that there is a waist on the almost triangular cigar. This is also consistent with the de Haas–van Alphen data. The larger cross-sectional area is about 2% greater than that of the waist. The corresponding experimental results predict 3%.

ACKNOWLEDGMENTS

The author is extremely grateful for the computer time made available at the Computation Center of the Pennsylvania State University.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Specific Heat of Thulium Metal Between 0.38 and 3.9°K*

O. V. LOUNASMAA†
Argonne National Laboratory, Argonne, Illinois
(Received 20 January 1964)

The specific heat C_p of thulium metal has been measured in a He^3 cryostat. Between 0.38 and 3.9°K $C_p=2.839T^3+17.94T+23.43T^{-2}-1.79T^{-3}-0.066T^{-4}$ (in mJ/mole °K). The last three terms represent the nuclear specific heat C_N. On the basis of earlier estimates, we put $C_L=0.243T^3$ and $C_E=10.5T$ for the lattice and electronic specific heats, respectively. According to the simple spin-wave theory, the magnetic specific heat C_M is proportional to T^3 for a ferrimagnetic metal; experimentally one finds $C_M=6.2T^{5/2}$ for thulium, which has a rather complicated ferrimagnetic structure. Further, there seems to be no evidence in C_M for an exponential factor, to be expected because of magnetic anisotropy. All conclusions on C_M are tentative, however, until data at temperatures between 4 and 20°K become available. C_N does not fit to the simple picture as given by Bleaney either. Since $I=\frac{1}{2}$ for the only stable thulium isotope Tm^{169}, quadrupole interactions are zero and there are only two nuclear energy levels, their separation being determined by the magnetic hyperfine constant a'. This would give a nuclear specific heat with even powers of T only, with a' determining the values of the coefficients. The observed C_N cannot be fitted into an equation of this type which indicates that other interactions, probably nuclear exchange interactions, are present. Formally, the experimental situation may be expressed by writing $a'=a_0-b/T$, instead of treating a' as a constant. Our results are in good agreement with recent Mössbauer data by Kalvius *et al.* who found 22.9 for the coefficient of the T^{-2} term.

I. INTRODUCTION

At low temperatures, the specific heat of rare-earth metals has four components which, depending on circumstances, can be separated totally or partially from each other. These are the lattice specific heat $C_L=AT^3$, the electronic specific heat $C_E=BT$, the magnetic specific heat C_M, and the nuclear specific heat C_N. In the higher lanthanides, C_M is primarily caused by exchange interaction between the $4f$ electronic spins. At 4.2°K and below, thulium has a unique ferrimagnetic structure, to be described in some detail later (cf. Sec. IV.B).[1] It is interesting to see how well the magnetic specific heat follows the prediction, $C_M=CT^3$, of the simple spin-wave theory.[2]

C_N is due to splitting of the nuclear spin states by interaction with the $4f$ electrons. By far the largest contribution to C_N comes from the magnetic field produced by the orbital angular momentum of the $4f$ electrons. The resulting nuclear specific heat has the familiar appearance of a Schottky curve with its maximum somewhere below 0.1°K. Above the maximum, C_N may be expressed in inverse powers of T, the leading term being proportional to T^{-2}.

According to the above discussion the total specific heat of thulium becomes

$$C_p=AT^3+BT+C_M+C_N, \qquad (1)$$

where A and B are constants.

The present measurements on thulium are a part of our research program for studying the heat capacities of rare-earth metals between 0.4 and 4°K,[3–10] with particular emphasis on C_N. The specific heat of thulium has previously been measured by Jennings, Hill, and

* Based on work performed under the auspices of the U. S. Atomic Energy Commission, and the Finnish Research Council for Technical Sciences.

† Present address: Wihuri Physical Laboratory, University of Turku, Turku, Finland.

[1] W. C. Koehler, J. W. Cable, E. O. Wollan, and M. K. Wilkinson, Phys. Rev. 126, 1672 (1962).

[2] J. van Kranendonk and J. H. van Vleck, Rev. Mod. Phys. 30, 1 (1958).

[3] O. V. Lounasmaa, Phys. Rev. 126, 1352 (1962).

[4] O. V. Lounasmaa and R. A. Guenther, Phys. Rev. 126, 1357 (1962).

[5] O. V. Lounasmaa and P. R. Roach, Phys. Rev. 128, 622 (1962).

[6] O. V. Lounasmaa, Phys. Rev. 128, 1136 (1962).

[7] O. V. Lounasmaa, Phys. Rev. 129, 2460 (1963).

[8] O. V. Lounasmaa, Phys. Rev. 133, A219 (1964).

[9] O. V. Lounasmaa, Phys. Rev. 133, A211 (1964).

[10] O. V. Lounasmaa, Phys. Rev. 133, A502 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Spedding[11] between 15 and 360°K and by Dreyfus, Goodman, Lacaze, and Trolliet[12] between 0.5 and 4°K. The former data do not extend to low enough temperatures for determining C_N. Measurements by Dreyfus *et al.* are published as a brief research note only and detailed comparisons are thus impossible; this is unfortunate in view of the interesting behavior observed for C_N in the present research (cf. Sec. IV.C).

II. EXPERIMENTAL

The heat capacity measurements were carried out in a He^3 cryostat which has been described earlier[4]; only the most important experimental features are related here. He^4 exchange gas was used for cooling the sample down to 4.2°K; the space surrounding the calorimeter was then evacuated by pumping until a mass-spectrometer-type leak detector showed a very small helium reading. For further cooling a mechanical heat switch was employed. In this way, good thermal insulation was achieved when the heat switch was opened and desorption of helium gas from the sample during heat capacity measurements was prevented. By pumping on He^3, a temperature of about 0.33°K was reached and maintained in the He^3 pot for 48 h without recondensing. With the heat switch closed, the sample was cooled from 4.2 to 0.35°K in about 24 h.

For the heat capacity measurements, a colloidal graphite (Aquadag) thermometer was employed; its construction has been described elsewhere.[4,8] The thermometer was calibrated against the vapor pressure of He^4 between 4.15 and 2.2°K, against the vapor pressure of He^3 between 2.2 and 0.75°K, and against a magnetic thermometer (chrominum methylamine alum salt) between 0.75 and 0.35°K. He^4 temperatures were determined according to the T_{58} scale.[13] For He^3 the 1962 Los Alamos scale[14] was employed; a correction was made for the 0.6% of He^4 in our He^3 gas. The constants in the susceptibility versus temperature curve for the magnetic thermometer were determined from calibration points between 2.2 and 0.75°K. Altogether about 30 points were measured for the carbon thermometer between 4.15 and 0.35°K.

III. RESULTS

Our thulium metal was purchased from Research Chemicals, Inc. (Division of Nuclear Corporation of America). It was vacuum distilled by the manufacturer, then remelted in a vacuum and cast into a tantalum crucible. Next, the tantalum was machined off and the sample turned down to a cylinder 0.9 cm long and 2.7 cm in diameter; its weight was 49.902 g (=0.29539 mole). The spectrographic laboratory at Argonne found the following metallic impurities (weight %): Al, 0.03%; Cr, 0.003%; Fe, 0.01%; K, 0.01%; Mg, 0.005%; Mn, 0.01%; Na, 0.02%, Ta, 0.12%; Ti, 0.01%. These analyses are uncertain by a factor of 2 except in the case of Ta for which the precision is ±10%. The Ta result was obtained by a very careful densitometric analysis. Of nonmetallic impurities, the Argonne chemical laboratory found: H, 0.0006%; N, 0.20%; O, 0.10%; F, 0.024%; C, 0.014%. The accuracy is about 10%.

The experimental specific heat results are listed in Table I and the points are also plotted into Fig. 1. All calculations were performed by an IBM-704 digital computer.[15] The results have been corrected for curvature due to finite temperature increments when measuring C_p. The size of these increments may be computed from the separation of successive points in Table I.

TABLE I. Specific heat (in mJ/mole °K) of thulium metal. Experimental results.

T(°K)	C_p	T(°K)	C_p	T(°K)	C_p
Run I		2.6331	101.5	Run III	
0.3876	129.6	2.8513	119.7	1.0345	42.04
0.4304	110.1	3.0789	140.9	1.1744	41.92
0.4808	92.70	3.3205	167.8	1.3213	43.54
0.5405	78.06	3.5976	199.1	1.4803	46.24
0.6114	65.71	3.9122	241.7	1.6372	50.29
0.6949	56.08			1.7819	54.75
0.7916	48.96	Run II		1.9211	60.12
0.8996	44.46	0.3799	133.6	2.0550	66.33
1.0157	42.26	0.4155	116.2	2.1880	72.38
1.1468	41.85	0.4572	100.2	2.3306	80.92
1.2877	43.24	0.5111	84.73	2.4937	91.61
1.4233	45.37	0.5746	71.73	2.6947	106.8
1.5598	48.40	0.6495	60.86	2.9315	126.8
1.6952	52.48	0.7359	52.63	3.1984	151.9
1.8286	57.04	0.8341	46.87	3.4993	187.5
1.9666	62.44	0.9428	43.44	3.8202	226.6
2.1092	69.19	1.0690	41.96		
2.2622	77.53	1.2077	42.49		
2.4347	88.37	1.3524	44.19		

After a heating period (about 1 min), the sample came to equilibrium in less than 2 sec at all temperatures. The heat leak to the calorimeter was always less than 0.2% of the corresponding heat input during heating periods. A correction for this was made in the customary manner by assuming linear drifts. The scatter of the points belonging to the same run from a smooth curve is about 0.3% (cf. Fig. 2). The heating current was measured with a Rubicon No. 2781 potentiometer and timed with an electronic timer using a tuning fork frequency standard. The potentiometer was cross checked against a Rubicon No. 2773 double potentiometer, the timer compared with radio signals over a 24-h period, and standard cells and standard resistors cali-

[11] L. D. Jennings, E. Hill, and F. H. Spedding, J. Chem. Phys. 34, 2082 (1961).
[12] B. Dreyfus, B. B. Goodman, A. Lacaze, and G. Trolliet, Compt. Rend. 253, 1764 (1961).
[13] F. G. Brickwedde, H. van Dijk, M. Durieux, J. R. Clement, and J. K. Logan, J. Research Natl. Bur. Std. A64, 1 (1960).
[14] R. H. Sherman, S. G. Sydoriak, and T. R. Roberts, Los Alamos Scientific Laboratory Report No. 2701 (unpublished).
[15] P. R. Roach, Argonne National Laboratory Technical Report No. 6497 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

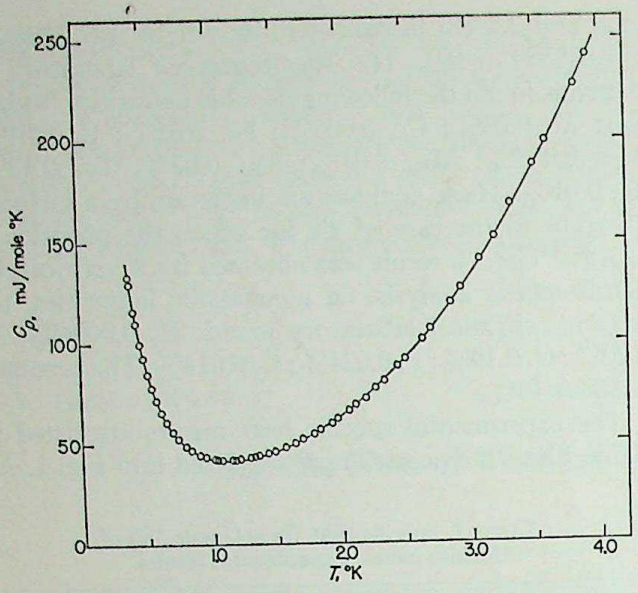


FIG. 1. The specific heat of thulium metal. Solid curve corresponds to equation $C_p=2.839T^3+17.94T+23.43T^{-2}-1.79T^{-3}-0.066T^{-4}$ (C_p in mJ/mole °K).

brated against NBS certified equipment. Possible systematic errors here are thus negligible. The electrical leads between the He^4 bath and the calorimeter were made of lead-covered constantan and were superconducting below 7°K. As described in another paper,[8] it is likely that the coating had tiny cracks at which a small amount of heat was generated. Due to this, there might be an error of 0.1% in the effective heater resistance ($R_H=362.13+0.03T$ Ω).

The heat capacity of the empty calorimeter ($C_{cal}=0.0090T^3+0.116T$ mJ/°K) was known from an earlier experiment; it was never more than 1.2% of the heat capacity of the sample. Possible uncertainties here can cause only errors smaller than 0.1% in the final results.

As is usual in low-temperature calorimetry, the largest errors in C_p are probably due to calibration of the carbon thermometer. The scatter of the calibration points was, with a few exceptions, less than 1 mdeg. The calculated temperature is thus probably within 1 mdeg of the temperature defined by the He^4 and He^3 scales.[13,14] Uncertainties of about 2 mdeg may arise from the extrapolation of the magnetic thermometer calibration below 0.75°K.

The three runs and the thermomenter calibrations were made during the same experiment without warming the sample above 4.2°K in the meantime. Largest discrepancies between runs occur in C_p at 2.0°K where they are about 0.5%.

After considering all the sources of errors and the discrepancies mentioned above, the accuracy of the present results is estimated as 0.6% between 1 and 4°K and 2% at 0.4°K.

The observed specific heat points from Table I, over the entire temperature range from 0.38 to 3.9°K, were fitted, after assuming $C_N=DT^{-2}-FT^{-4}(F=D^2/R$, cf. Sec. IV.C) and $C_M=CT^3$, to Eq. (1) by the method of least squares. The result was (C_p in mJ/mole °K):

$$C_p=2.74T^3+19.2T+19.15T^{-2}-0.044T^{-4}. \quad (2)$$

However, as may be seen from Fig. 2, deviations of the measured points from this curve are fairly large at low temperatures. This indicates that a term proportional to T^{-3} must be taken into account when expressing C_N (for a discussion on the temperature dependence of C_N, see Sec. IV.C). By writing $C_N=DT^{-2}-ET^{-3}-FT^{-4}$ the least-squares result was

$$C_p=2.839T^3+17.94T+23.43T^{-2}-1.79T^{-3}-0.066T^{-4}. \quad (3)$$

Standard deviations of the first four coefficients are 0.012, 0.12, 0.26, and 0.11. Equation (3) gives a reasonably good fit to the experimental points over the entire temperature range from 0.38 to 3.9°K (cf. Fig. 1); maximum deviations in $\Delta C/C$ are 1.9% at 1.3°K. Figure 2 shows that the fit is much better than with Eq. (2).

IV. DISCUSSION

A. The Lattice and Electronic Heat Capacities

Since the temperature dependence of C_M is uncertain, we must assume that the first term in (3) is due to C_L and C_M, the second term to C_E and C_M, and the last three terms to C_N. It is thus impossible to separate C_L and C_E from C_M without additional information.

On the basis of considerations described in an earlier paper,[8] we estimate that the lattice heat capacity of thulium should correspond to a Debye characteristic temperature $\theta=200$°K giving $C_L=0.243T^3$, and that $C_E=10.5T$ (C_L and C_E in mJ/mole °K). The major contribution to the first term in Eq. (3) thus comes

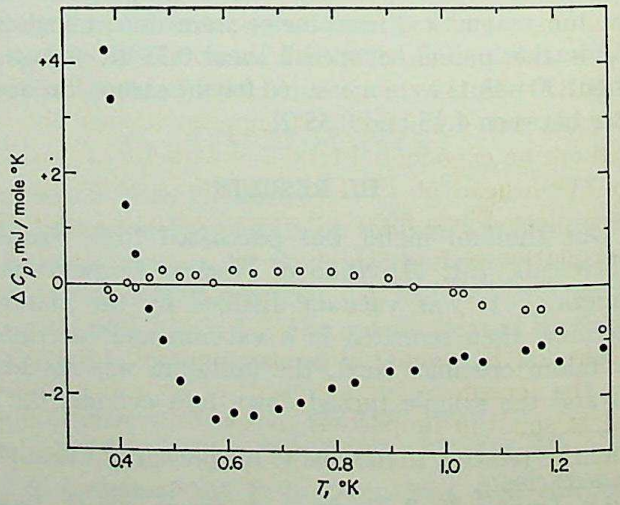


FIG. 2. Deviations ΔC_p of the experimental specific heat points from the expressions $C_p=2.74T^3+19.2T+19.15T^{-2}-0.044T^{-4}$ [filled circles, Eq. (2)] and $C_p=2.839T^3+17.94T+23.43T^{-2}-1.79T^{-3}-0.066T^{-4}$ [open circles, Eq. (3)]. $\Delta C_p=C_p$(calc.) $-C_p$(exp) (in mJ/mole °K).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

from C_M. The precision of these values is estimated as 5% for θ and 7% for C_E.[8]

B. The Magnetic Heat Capacity

The magnetic susceptibility of polycrystalline thulium has been studied by Rhodes, Legvold, and Spedding[16] and by Davis and Bozorth.[17] A Néel point was observed at about 56°K and a Curie point at 22°K. According to Jennings, Hill, and Spedding[11] the heat capacity shows a pronounced peak at the Néel point but no anomalous behavior was noted at the Curie point.

The magnetic structure of thulium has been determined by Koehler, Cable, Wollan, and Wilkinson[1] by using neutron diffraction techniques. At 4.2°K, where saturation has very nearly been reached, thulium has an antiphase domain-type structure in which four layers of atoms with their magnetic moments pointing in the direction of the positive c axis are followed by three layers of atoms with their moments in the opposite direction. The low-temperature magnetic structure of thulium is thus ferrimagnetic. Each atom has an ordered magnetic moment of 7 μ_B, and the ferrimagnetic structure has thus a net moment of 1 μ_B per atom.

According to the simple spin-wave theory,[2] the magnetic specific heat of a ferrimagnetic metal, at temperatures well below the Curie point, may be written $C_M=CT^3$, where C is a constant. Niira[18] and Cooper[19] have shown, however, that in metals with magnetic anisotropy there is an energy gap E_g at the bottom of the spin-wave spectrum. Due to this, $C_M=CT^{3/2}\exp(-E_g/kT)$ for a ferromagnetic metal and $C_M=CT^3\exp(-E_g/kT)$ for a ferri- or antiferromagnetic metal.[20] More complicated formulas have also been proposed.[19] The physical reason for E_g is that it always takes a finite energy to turn a spin against the anisotropy field. Energy gaps of the order of $E_g/k=20-30$°K have been predicted for dysprosium[19] and for terbium.[5]

We have tried to fit the magnetic specific heat of thulium, calculated from the relation (in mJ/mole °K; C_p is the measured specific heat) $C_M=C_p-0.243T^3-10.5T-23.43T^{-2}+1.79T^{-3}+0.066T^{-4}$, to the above formula for C_M. The best agreement was obtained without an exponential term and with C_M proportional to $T^{5/2}$ instead of T^3. The magnetic specific heat has been plotted into Fig. 3; the straingth line corresponds to equation (C_M in mJ/mole °K).

$$C_M=6.2T^{5/2}. \tag{4}$$

Above 2.2°K, where C_M is more than 60% of the total C_p, the agreement is quite good; at lower temperatures C_M is small in proportion to the other contributions and the relative precision with which it can be determined becomes poor. Of course, the weak points in our calculation are the assumptions made about the size of C_L and C_E. By writing $C_E=17.94T$ [cf. Eq. (3)], C_M would be proportional to T^3. It seems unlikely, however, that the electronic specific heat is much larger for thulium than for the other trivalent lanthanides.[8] Of course, by subtracting C_L, C_E, and C_N from C_p in Eq. (3), we get $C_M=2.596T^3+7.44T$. This relation for C_M has, however, no theoretical basis.

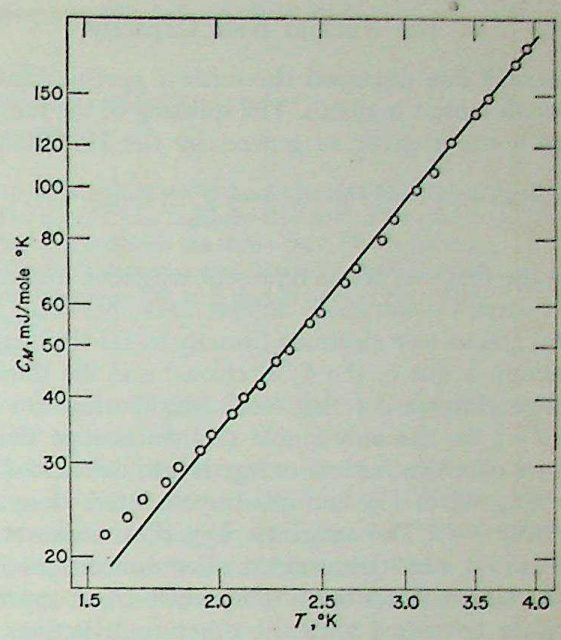


FIG. 3. The magnetic specific heat of thulium. The straight line corresponds to equation $C_M=6.2T^{5/2}$ (C_M in mJ/mole °K).

Before drawing any definite conclusions about the temperature dependence of C_M, measurements of C_p should be made at higher temperatures, especially between 4 and 20°K. This would allow us to decide whether or not an exponential factor is necessary in expressing C_M. At the moment one can only say that such a factor does not seem to be very important.

Measurements by Dreyfus *et al.*[12] gave $C_p=0.949T^3+21.5T+27T^{-2}$ (C_p in mJ/mole °K). Their data agree with ours at 1.5°K, are higher at lower temperatures and lower above 1.5°K; at 4°K the difference is 40%. Large discrepencies between specific heat results of different investigators are quite common between 2 and 4°K for rare-earth metals.[3–5,9] The disagreement is usually attributed to impurities which, even in small concentrations, might significantly alter the energy gap E_g at the bottom of the spin-wave spectrum.[4,5,19] There is some difficulty with this explanation here since the exponential factor in C_M seems to be absent or its effect is small in thulium. Another possibility is magnetic ordering in an impurity which, even in quantities of 0.1% of the sample, will cause large contributions to the observed specific heat, provided a magnetic transformation occurs in the impurity in or near the temperature range under investigation.[7]

[16] B. L. Rhodes, S. Legvold, and F. H. Spedding, Phys. Rev. **109**, 1547 (1958).
[17] D. D. Davis and R. M. Bozorth, Phys. Rev. **118**, 1543 (1960).
[18] K. Niira, Phys. Rev. **117**, 129 (1960).
[19] B. R. Cooper, Proc. Phys. Soc. (London) **80**, 1225 (1962).
[20] A. R. Mackintosh, Phys. Letters **4**. 140 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

C. The Nuclear Heat Capacity

Bleaney[21] has discussed the nuclear specific heat of rare-earth metals in detail. The splitting of the nuclear energy levels is given in general by the Hamiltonian

$$\mathcal{H}=a'i+P[i^2-\tfrac{1}{3}I(I+1)]-\tfrac{1}{2}\sum_{j\neq k}2(\mathbf{I}_j\cdot\mathbf{K}_{jk}\cdot\mathbf{I}_k), \quad (5)$$

where the first two terms represent magnetic hyperfine and electric quadrupole interactions between the nucleus and its own electrons (usually by far the biggest interaction is due to the $4f$ electrons) and the third is exchange interaction between neighboring nuclei. Since $I=\frac{1}{2}$ for the only stable thulium isotope Tm^{169}, there are only two nuclear energy levels, corresponding to $i=+\frac{1}{2}$, $-\frac{1}{2}$ in (5), and quadrupole interactions are identically $=0$. The magnetic hyperfine constant a' varies as $\langle J_z\rangle$, which measures the electronic magnetization. At temperatures below 2°K, where C_N is appreciable when compared with the other contributions to C_p, complete electronic magnetization is usually assumed and $\langle J_z\rangle$ is replaced by $J(=6$ for Tm^{3+} ions) and thus $a'=$const.

Once the nuclear energy levels are known, C_N can be calculated from the partition function in the usual way and, by expanding the nuclear specific heat in inverse powers of T, we obtain [cf. Eq. (3)]

$$C_N=DT^{-2}-ET^{-3}-FT^{-4}+\cdots, \quad (6)$$

where

$$D/R=\tfrac{1}{3}(a')^2I(I+1)+(4/9)I^2(I+1)^2\sum_{jk}(K_x^2+K_y^2+K_z^2)_{jk}, \quad (7)$$

$$E/R=(4/9)(a')^2\sum_{jk}(K_{jk})_zI^2(I+1)^2, \quad (8)$$

$$F/R=(1/30)(a')^4I(I+1)(2I^2+2I+1) \quad (9)$$

($P=0$, R is the gas constant, magnetic field is assumed in z direction, some higher terms in $\mathbf{K}$ have been ignored).

According to Bleaney's[21] estimates, contributions from the nuclear exchange interaction should be relatively small. However, if we put $\mathbf{K}=0$ only even powers of T appear in (6) and ($I=\frac{1}{2}$)

$$F=(3D^2/10R)(2I^2+2I+1)/(I^2+I)=D^2/R. \quad (10)$$

By assuming $D=19.15$ mJ °K/mole [cf. Eq. (2)], we get $F=0.044$ mJ °K³/mole. Even at 0.4°K, the T^{-4} term accounts only for 1.5% of the total C_N. By ignoring relation (10) and by letting all the coefficients in (2) be free in the least-squares analysis, we obtain a considerably larger value of F (C_p in mJ/mole °K):

$$C_p=2.82T^3+18.2T+21.6T^{-2}-0.49T^{-4}. \quad (11)$$

This equation fits to the measured points almost as well as (3). Our experimental accuracy is not sufficient to decide whether the "residual" C_N is proportional to T^{-3} or to T^{-4}.

In any case, the observed C_p below 1°K forces us to assume that other interactions, besides the magnetic hyperfine term, have a fairly large effect in the value of C_N. Due to the peculiar magnetic structure of thulium, the nuclear exchange interaction, acting through the spin waves, might be unusually large in this metal. We shall adopt Eq. (3) for representing the specific heat of thulium; this choice was made since it is impossible to explain the large T^{-4} term in (11) on theoretical grounds. In these two relations the coefficient of the T^{-2} term differs by 8%. On the basis of this and our earlier error estimates (cf. Sec. III), we put $D=23.4\pm1$ mJ °K/mole. If the temperature dependence of C_N in Eq. (3) and the ratio $F/D=D/R$ [cf. Eq. (10)] can be assumed correct, the limits of error in D would be ±0.5 mJ °K/mole.

Dreyfus *et al.*[12] have calculated from their specific heat data $D=27\pm3$ mJ °K/mole without taking higher terms in C_N into account. Since their measurements have been published as a smoothed curve only, it is impossible to reanalyze them. However, inclusion of higher terms would tend to increase D. From atomic beam experiments on neutral thulium atoms by Ritter,[22] Bleaney[21] has derived $D=26.9$ mJ °K/mole, which is in good agreement with the result by Dreyfus *et al.* However, recent Mössbauer measurements by Kalvius, Kienle, Eicher, Wiedemann, and Schüler[23] on thulium metal gave, at 5°K, the splitting of the ground state as $9.04\cdot10^{-6}$ eV, which corresponds to $D=22.9$ mJ °K/mole. If electronic magnetization was not complete at 5°K, this value would tend to increase towards lower temperatures; this effect cannot be larger than about 0.5%.[23]

The present specific heat measurements are thus in good agreement with the Mössbauer data.[23] The observed temperature dependence of C_N remains a puzzle. Formally one can write for thulium, in the temperature range studied, $a'=a_0-b/T$, whatever the reason may be. Here a_0 and b are constants. By extending the measurements to still lower temperatures, more would be learned about the nature of the interactions, in particular, one could see whether cooperative effects are present.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor B. Bleaney and Dr. R. J. Elliott, Dr. R. W. Hill, and Dr. A. R. Mackintosh for useful comments and to the chemical and spectrochemical laboratories at Argonne for analyses of the samples. He is also indebted to Osk. Huttusen Säätiö for a research grant.

[21] B. Bleaney, J. Appl. Phys. 34, 1024 (1963).

[22] G. J. Ritter, Phys. Rev. 128, 2238 (1962).

[23] M. Kalvius, P. Kienle, H. Eicher, W. Wiedemann, and C. Schüler, Z. Physik 172, 231 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Collective Strain Splitting of Acceptor States in Silicon*

CALVIN S. KOONCE†
Department of Physics, University of California, Berkeley, California
(Received 27 January 1964)

A splitting of the ground acceptor state in silicon due to equilibrium homogeneous strains and a corresponding transition temperature have been calculated. The equilibrium homogeneous strains are due to linear terms in the strain Hamiltonian which couple the acceptor states with the strain field. The transition temperature was found to be $N_A \times 5.6 \times 10^{-20}$ °K, where N_A is the acceptor concentration in acceptors per cubic centimeter. For indium, concentrations as high as 10^{20} acceptors per cc should be possible before Coulomb and exchange effects between acceptors are dominant.

I. INTRODUCTION

THE acceptor state in silicon having the lowest energy transforms like Γ_8 of the tetrahedral double group. That is, the lowest acceptor state possesses the same fourfold degeneracy as the $P_{3/2}$ valence band edge at $k=0$.[1] At temperatures sufficiently low, all holes can be considered to be in the ground acceptor state.

It has been found[2,3] that the valence band edge of silicon and germanium at $k=0$ is split under the application of a homogeneous uniaxial stress. These stresses also split the degeneracy of the ground acceptor state.[4]

If there are two acceptor states in the crystal, they should be able to interact through the strain, or phonon, field. If there are a great many acceptor states, they should collectively produce a strain in a manner analogous to the way in which a lattice of magnetic dipoles interact to produce ferromagnetism.

In this paper the lowering of the energy corresponding to this collective equilibrium strain and a corresponding transition temperature are calculated.

II. DIAGONALIZATION OF THE HAMILTONIAN

In this section a Hamiltonian density will be set up and diagonalized, yielding the energy density associated with various equilibrium strains in the crystal. The strains associated with the minimum energy density and the maximum lowering of the energy density are obtained.

Although the strain Hamiltonian for silicon doped with acceptor atoms contains strains of all wavelengths, and the equilibrium strain in the neighborhood of an acceptor state will contain strains having wavelengths comparable to the dimensions of the acceptor state and even shorter, the strains responsible for the collective effect of all acceptor states are homogeneous strains. Therefore, in all that follows, only homogeneous strains will be considered.

If an individual acceptor is considered, it can strain in one of the normal-mode strains. However, the zero-point motion is large enough to average over all of the possible normal-mode strains and their negatives, so that on the average no strain exists. The only effect is to increase the amplitude of the zero-point motion. A homogeneous strain, however, induces a splitting even after averaging over the zero-point motions. Only the states split by the homogeneous strain are considered to be strain states of the system. This is in contrast with treatments of Jahn-Teller phase transformations of ordered impurities in which the individual atomic strain states are considered.[5]

In the strain Hamiltonian there are two types of terms. First, there are the terms independent of strain, quadratic in the strain and of higher order in the strain which are present when the acceptor concentration, and hence the hole concentration, is zero. All powers of the strain higher than the second will be neglected.

Second, there are terms which depend on the acceptor concentration. Because the acceptor concentration will always be much less than the atomic concentration of silicon, strain terms quadratic and higher in the strain which depend on acceptor concentration will be neglected. The change in elastic constants with acceptor concentration has been computed only for quite high acceptor concentrations[6] and will be neglected here. The linear terms in the Hamiltonian exist because the acceptor state is a wave packet of states taken from the top of the valence band, and exhibits the degeneracy in the valence band. The Jahn-Teller effect[7] states that, if a nonlinear symmetric arrangement of atoms has a degenerate ground state (not including spin degeneracy), there will exist in the Hamiltonian nonvanishing terms linear in the displacement of the atoms which destroy the symmetry and split the degeneracy.

The most general form for these linear terms for a crystal with cubic symmetry has been given by Kleiner

* Supported by the National Science Foundation.
† National Science Foundation Predoctoral Fellow.
[1] W. Kohn, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5.
[2] W. H. Kleiner and C. M. Roth, Phys. Rev. Letters 2, 334 (1959).
[3] J. C. Hensel and G. Feher, Phys. Rev. 129, 1041 (1963).
[4] G. Feher, J. C. Hensel, and E. A. Gere, Phys. Rev. Letters 5, 309 (1960).
[5] P. J. Wojtowicz, Phys. Rev. 116, 32 (1959).
[6] P. Csavinszky and N. G. Einspruch, Phys. Rev. 132, 2435 (1963).
[7] H. A. Jahn and E. Teller, Proc. Roy. Soc. (London) 161, 220 (1937).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

and Roth[2] and can be obtained from symmetry considerations alone. The values of the deformation potential coefficients D_u and D_u' for states in the valence-band edge of silicon at $k=0$ have been determined by Hensel and Feher.[3] Because the ground acceptor state has the same symmetry as the valence-band edge, the linear terms caused by the acceptor state will have exactly the same form as those for the valence-band edge, except for a change in the deformation potential coefficients.

If the binding energy of the acceptor state were much less than the energy lowering due to the equilibrium strain, the deformation potential coefficients obtained for the valence band edge by Hensel and Feher could be used for the acceptor state. Although this condition is not satisfied, these deformation potential coefficients will be used. The Hamiltonian density due to acceptor states is taken to be the acceptor concentration times the energy per hole given by Kleiner and Roth.[2]

The Hamiltonian density is then

$$\begin{aligned}\mathcal{H}=\mathcal{H}_0+(\tfrac{1}{2}c_{11})\sum_i e_{ii}^2+c_{12}\sum_{i>j}e_{ii}e_{jj}+(\tfrac{1}{2}c_{44})\sum_{i>j}e_{ij}^2\\ +N_A\{D_d^v\sum_i e_{ii}+(\tfrac{2}{3}D_u)\sum_i(J_i^2-\tfrac{1}{2}J^2)e_{ii}\\ +(\tfrac{2}{3}D_u')\sum_{i>j}[(J_iJ_j+J_jJ_i)/2]e_{ij}\}\,,\end{aligned}\quad(1)$$

where c_{ii} and c_{ij} are the elastic stiffness constants, e_{ii} and e_{ij} are the conventional strain components,[3] D_d^v, D_u, and D_u' are the deformation potential coefficients appropriate to the various types of strain, $\mathcal{H}_0$ is the Hamiltonian density with no strain, N_A is the acceptor concentration, J_i is an angular momentum matrix for angular momentum $\frac{3}{2}$, and the sums over i and j run from one to three over the Cartesian components. Where no angular momentum matrix is indicated, the identity matrix should be taken, because the terms independent of acceptor concentration do not split the acceptor states.

If a conventional phase for our angular momentum matrices is chosen,[8] we can write our Hamiltonian in matrix form.

$$\mathcal{H}=\begin{pmatrix}\epsilon+Q_1+Q_2 & Q_5-iQ_6 & Q_3-iQ_4 & 0\\ Q_5+iQ_6 & \epsilon+Q_1-Q_2 & 0 & Q_3-iQ_4\\ Q_3+iQ_4 & 0 & \epsilon+Q_1-Q_2 & -Q_5+iQ_6\\ 0 & Q_3+iQ_4 & -Q_5-iQ_6 & \epsilon+Q_1+Q_2\end{pmatrix},\quad(2)$$

where

$$\begin{aligned}Q_1&=N_AD_d^v(e_{xx}+e_{yy}+e_{zz});\\ Q_2&=N_A(\tfrac{2}{3}D_\mu)[e_{zz}-\tfrac{1}{2}(e_{xx}+e_{yy})];\\ Q_3&=N_A(\tfrac{2}{3}D_u)(\tfrac{1}{2}\sqrt{3})(e_{xx}-e_{yy});\\ Q_4&=N_A(\tfrac{2}{3}D_u')(\tfrac{1}{2}\sqrt{3})e_{xy};\\ Q_5&=N_A(\tfrac{2}{3}D_u')(\tfrac{1}{2}\sqrt{3})e_{xz};\\ Q_6&=N_A(\tfrac{2}{3}D_u')(\tfrac{1}{2}\sqrt{3})e_{yz},\end{aligned}\quad(3)$$

and

$$\epsilon=E_0+(\tfrac{1}{2}c_{11})\sum_i e_{ii}^2+c_{12}\sum_{i>j}e_{ii}e_{jj}+(\tfrac{1}{2}c_{44})\sum_{i>j}e_{ij}^2;\quad(4)$$

or, in terms of the Q_i,

$$\begin{aligned}\epsilon=E_0+\frac{c_{11}+2c_{12}}{6(D_d^vN_A)^2}Q_1^2+\frac{c_{11}-c_{12}}{3(\frac{2}{3}N_AD_u)^2}(Q_2^2+Q_3^2)\\ +\frac{2c_{44}}{3(\frac{2}{3}N_AD_u')^2}(Q_4^2+Q_5^2+Q_6^2).\end{aligned}\quad(5)$$

This matrix can be diagonalized to give eigenvalues for the energy density,

$$\lambda=\epsilon+Q_1\pm(Q_2^2+Q_3^2+Q_4^2+Q_5^2+Q_6^2)^{1/2}.\quad(6)$$

The minimum for the energy density cannot depend on the particular choice of Q_2 and Q_3 as long as $Q_2^2+Q_3^2$ is equal to a constant. The same degeneracy holds for Q_4, Q_5, and Q_6. Also, strains of the type Q_1 are separated from the other types and cause no energy splitting. Minimizing the energy density λ with respect to Q_1, one obtains an equilibrium value for Q_1. Terms in Q_1 can be included with the terms independent of strain to form the arbitrary zero of our energy.

The energy density change due to the other strains is

$$\lambda'=Ax^2+By^2\pm(x^2+y^2)^{1/2},\quad(7)$$

where

$$\begin{aligned}A&=\tfrac{1}{3}(c_{11}-c_{12})(\tfrac{2}{3}N_AD_u)^2;\quad B=\tfrac{1}{3}2c_{44}(\tfrac{2}{3}N_AD_u')^2;\\ x^2&=Q_2^2+Q_3^2;\qquad y^2=Q_4^2+Q_5^2+Q_6^2.\end{aligned}\quad(8)$$

The energy density can be minimized with respect to x and y by finding the values for x and y for which the first derivative of λ' with respect to both x and y is equal to zero. For $A\neq B$, as is usually the case, either x or y must be zero at equilibrium. However, $x=y=0$ gives $\lambda'=0$ and is not a minimum for λ'. If $A>B$, as is true for silicon, the minimum energy density is

$$\lambda_{\min}'=-1/4B=-N_A^2D_u'^2/6c_{44},\quad(9)$$

and the corresponding distortions are

$$\begin{aligned}e_{xy}^2+e_{xz}^2+e_{yz}^2&=N_A^2D_\mu'^2/3c_{44}^2;\\ e_{xx}=e_{yy}&=e_{zz}=Q_1/3N_AD_d^v.\end{aligned}\quad(10)$$

The strains e_{xx}, e_{yy}, and e_{zz} are chosen so that Q_2 and Q_3 are zero, and their value is determined by the value of Q_1 at equilibrium.

III. THE WEISS FIELD METHOD

The analogy with ferromagnetism will now be pursued, letting the strain field play the part of the magnetic field, and the average value of the angular momentum components of the acceptor state play the part of the magnetization. It will be found that the ratio of the average angular momentum components

[8] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., p. 146.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

to the applied strains obeys a "Curie-Weiss law," and the transition temperature can be computed.

The response of the crystal to an applied strain must be determined. If the applied strain is of the form e_{xy}, then a "Weiss field" of the same form is assumed.

$$e_0=e_{xy}|_0;\quad e_{xx}|_0=e_{yy}|_0=e_{zz}|_0=e_{xz}|_0=e_{yz}|_0=0. \quad (11)$$

The notation $|_0$ means to evaluate the strain with no external strain fields applied.

$$e_{xy}=e_0+e_{xy}';\quad e_{xx}=e_{xx}';\quad \text{etc.}, \quad (12)$$

where e_{xy}' is the applied strain.

The Hamiltonian for the splitting of the acceptor state due to the total strain field e_{xy} is

$$\mathcal{H}_S=\tfrac{2}{3}D_\mu'[(J_xJ_y+J_yJ_x)/2]e_{xy}. \quad (13)$$

After diagonalizing (13), one finds that the acceptor level is split into two states, each twofold degenerate, with an energy difference

$$\Delta E=|(2/\sqrt{3})D_u'e_{xy}|. \quad (14)$$

For the higher energy state, the eigenfunctions are $(\psi_{1/2}+i\psi_{-3/2})/\sqrt{2}$; $(\psi_{3/2}+i\psi_{-1/2})/\sqrt{2}$. This state will be denoted by $+$. The lower energy state has eigenfunctions $(\psi_{1/2}-i\psi_{-3/2})\sqrt{2}$; $(\psi_{3/2}-i\psi_{-1/2})/\sqrt{2}$. This state will be denoted by $-$.

The density of states at each level is

$$N_+=\frac{N_A\exp(-D_u'e_{xy}/\sqrt{3}kT)}{\exp(D_u'e_{xy}/\sqrt{3}kT)+\exp(-D_u'e_{xy}/\sqrt{3}kT)};$$
$$N_-=\frac{N_A\exp(D_u'e_{xy}/\sqrt{3}kT)}{\exp(D_u'e_{xy}/\sqrt{3}kT)+\exp(-D_u'e_{xy}/\sqrt{3}kT)}. \quad (15)$$

The density of the angular momentum coefficient is

$$\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}=\tfrac{1}{2}\langle+|J_xJ_y+J_yJ_x|+\rangle N_+ +\tfrac{1}{2}\langle-|J_xJ_y+J_yJ_x|-\rangle N_-. \quad (16)$$

Here $\langle+||+\rangle$ indicates that the expectation value between the upper states should be taken. This gives the upper eigenvalue $+\sqrt{3}/2$.

Substituting (15) into (16), one obtains

$$\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}=-N_A(\sqrt{3}/2)\tanh(D_u'e_{xy}/\sqrt{3}kT), \quad (17)$$

This form of strain dependence has been experimentally verified for the case of O_2 centers in alkali halides.[9] Equation (17) becomes, for $D_u'e_{xy}\ll kT$,

$$\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}=-(N_AD_u'/2k)(e_{xy}/T)=C(e_{xy}/T). \quad (18)$$

The expectation values between $+$ or $-$ states for all other angular momentum combinations in our Hamiltonian (1) vanish, so that the assumed strain gives an average value for only its coefficient.

This average splitting of the acceptor states in turn leads to an equilibrium strain in the crystal. The energy density is

$$\mathcal{E}=\tfrac{2}{3}D_u'[\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}]e_{xy} +\tfrac{1}{2}c_{11}\sum_i e_{ii}^2+c_{12}\sum_{i>j}e_{ii}e_{jj} +\tfrac{1}{2}c_{44}\sum_{i>j}e_{ij}^2. \quad (19)$$

The energy density is then minimized with respect to the possible strains to obtain the equilibrium strains

$$e_{xx}|_0=e_{yy}|_0=e_{zz}|_0=e_{xz}|_0=e_{yz}|_0=0;$$
$$e_{xy}|_0=e_0=-\tfrac{2}{3}D_u'[\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}] =\gamma[\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}]. \quad (20)$$

Equations (12), (18), and (20) yield

$$\frac{\tfrac{1}{2}\langle J_xJ_y+J_yJ_x\rangle_{\rm av}}{e_{xy}'}=\frac{C}{T-C\gamma}. \quad (21)$$

$C\gamma$, the transition temperature, is given by

$$T_c=C\gamma=N_AD_u'^2/3kc_{44}. \quad (22)$$

Comparing (9) and (22), one obtains for the lowering of the energy density

$$\lambda_{\rm min}'=-\tfrac{1}{2}N_AkT_c. \quad (23)$$

The splitting of the lowest acceptor level is then

$$\Delta E=kT_c, \quad (24)$$

which might have been expected.

Using $D_u'=2.68\pm0.25$ eV, determined by Hensel and Feher,[3] and a value for $c_{44}=8.007\times10^{11}$ dyn/cm^2, which it has at low temperatures and which varies little with temperature at low temperatures,[6,10] one obtains

$$T_c=N_A\times5.6\times10^{-20}\text{°K}=n_A\times2800\text{°K}, \quad (25)$$

where n_A is the fractional number of acceptors and N_A is the acceptor concentration in acceptors per cc.

Feher, Hensel, and Gere[4] found that a stress of about 300 kg/cm^2 in the (100) direction was necessary to split the ground acceptor state enough for paramagnetic resonance absorption transitions to be observed over the broadening due to the background of random strains in the crystal. The energy splitting due to an applied stress in the [100] direction[3] is

$$\Delta E_S=\tfrac{4}{3}D_uT/(c_{11}-c_{12}), \quad (26)$$

where T is the applied stress. The splitting necessary for transitions to be observed is roughly 8×10^{-5} eV, or 0.9°K.

In order to observe the effect of the equilibrium strain, one would like to increase the concentration to a value which would allow the energy splitting of (24) to be greater than the energy splitting necessary to overcome the random strains in the crystal. However,

[9] W. Känzig, Phys. Chem. Solids 23, 479 (1962).

[10] H. J. McSkimin, J. Appl. Phys. 24, 988 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE I. Approximate concentrations and transition temperatures.

Acceptor	N_A in acc/cc	T_c in °K
B	10^{18}	0.06
Al	3×10^{18}	0.17
Ga	4×10^{18}	0.22
In	10^{20}	5.6

at very high concentrations, the acceptor states gain more energy by binding into hydrogen-like molecules and finally bands than they gain due to the Jahn-Teller effect. The energy of binding into hydrogen-like molecules comes from both direct Coulomb and exchange interactions. For concentrations in the range of interest, this binding energy is much larger than the energy due to Van der Waals forces between acceptor states and the direct magnetic coupling between acceptor states.

Therefore, one must know at what concentration the molecular binding energy exceeds the energy lowering computed from (9). Table I gives the approximate concentrations at which this occurs, and the transition temperatures associated with these concentrations.

ACKNOWLEDGMENT

The author would like to thank Professor J. J. Hopfield for proposing the problem and for many helpful suggestions and discussions.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Band Structure of Gallium Phosphide from Optical Experiments at High Pressure*

RICHARD ZALLEN AND WILLIAM PAUL
Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts

(Received 22 January 1964)

The effect of hydrostatic pressure on the following optical properties of GaP has been measured at room temperature: the fundamental absorption edge region from 2.2 to 2.7 eV, an infrared absorption band appearing in n-type material at 0.3–0.5 eV, peaks in the reflectivity spectrum at 2.8 and 3.7 eV, and recombination radiation in forward-biased p-n junctions at 1.7–2.3 eV. The results have been interpreted by means of a proposed energy band structure in which the conduction band states X_1^c, X_3^c, Γ_1^c, Γ_{15}^c are located at energies of 2.2, 2.5, 2.8, 3.7 eV, respectively, above the valence band maximum at Γ_{15}^v. The following pressure coefficients have been measured (the transition involved is given in parenthesis), where energy is expressed in eV and pressure in 10^6 bars: $E_G(\Gamma_{15}^v \rightarrow X_1^c) = 2.22 - 1.1P$; $E_0(\Gamma_{15}^v \rightarrow \Gamma_1^c) = 2.78 + 10.7P$; $E_0'(\Gamma_{15}^v \rightarrow \Gamma_{15}^c) = 3.71 + 5.8P$; $\Delta E_2(X_1^c \rightarrow X_3^c) = 0.3 + 1P$. The coefficients of E_G and E_0' are close to those for the corresponding transitions in Si; that of E_0 is close to the corresponding coefficient in Ge. The weak reflectivity peak at 2.8 eV, the direct gap, shifts with temperature at a rate of about -4.6×10^{-4} eV/°K, compared to a value of about -5.2×10^{-4} eV/°K for the 2.2 eV indirect gap.

1. INTRODUCTION

THIS paper describes an experimental investigation of the electronic energy band structure of GaP. The approach taken in this study was to measure the effect of hydrostatic pressure on certain optical properties in the infrared, visible, and near ultraviolet regions of the spectrum. Included in these were absorption near the fundamental absorption edge, extrinsic absorption, reflectivity, and recombination radiation. Several features of the lowest conduction bands have been elucidated in the energy range from 2 to 4 eV above the valence band maximum.

Part of the basis for this work is a systematic correlation inferred from earlier work on pressure effects in group 4 and group 3–5 semiconductors.[1–8] It has been observed that corresponding energy separations in these materials have similar pressure coefficients. In other words, the effect of pressure on the energy of an electronic transition depends primarily on the particular type of final and initial states involved (each specified by a band index n, a reduced wave vector k, and a group-theoretic classification), and is relatively insensitive to choice of material from among this class of simple semiconductors. We have listed some pressure

* Research supported by the U. S. Navy Office of Naval Research.

[1] Some of the principal results of this work were briefly mentioned in R. Zallen and W. Paul, Bull. Am. Phys. Soc. 9, 61 (1964).

[2] W. Paul and D. M. Warschauer, Phys. Chem. Solids 5, 89 (1958); 5, 102 (1958); 6, 6 (1958).

[3] W. Paul, Phys. Chem. Solids 8, 196 (1959).

[4] W. Paul, J. Appl. Phys. 32, 2082 (1961).

[5] W. Paul and D. M. Warschauer, *Solids under Pressure* (McGraw-Hill Book Company, Inc., New York, 1963), Table 8–4, p. 226.

[6] W. Paul and H. Brooks, Progr. Semiconductors 7, 135 (1963).

[7] R. Zallen, W. Paul, and J. Tauc, Bull. Am. Phys. Soc. 7, 185 (1962); some of these data are shown in the article by J. Tauc, in *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 341.

[8] R. Zallen, Gordon McKay Laboratory of Applied Science, Harvard University, Technical Report HP-12, 1964 (unpublished); R. Zallen and W. Paul, unpublished data.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

coefficients in Table I and briefly discuss two examples below.

TABLE I. Pressure coefficients of some energy separations in group 4 and group 3–5 semiconductors.

Transition (low energy→high energy) group 4 element	group 3–5 compound	$\partial E/\partial P$ in units of 10^{-6} eV/bar	Representative materials measured
$\Gamma_{25'} \rightarrow \Gamma_{2'}$	$\Gamma_{15} \rightarrow \Gamma_1$	$+(10-15)$	Ge,[a] GaAs,[b] InSb[c]
$\Gamma_{25'} \rightarrow L_1$	$\Gamma_{15} \rightarrow L_1$	$+5$	Ge,[d] Sn[e]
$\Gamma_{25'} \rightarrow \Delta_1$	$\Gamma_{15} \rightarrow \Delta_1$	$-(1-2)$	Si,[f] Ge[g]
$L_{3'} \rightarrow L_1$	$L_3 \rightarrow L_1$	$+(7-8)$	Ge, GaSb, InAs[h]

[a] See Ref. 10.
[b] See Ref. 11.
[c] See Ref. 12.
[d] See Ref. 3.
[e] S. H. Groves, Gordon McKay Laboratory of Applied Science, Harvard University, Technical Report HP-10, 1963 (unpublished.)
[f] See Refs. 2 and 26.
[g] See Refs. 3 and 38.
[h] See Ref. 7.

(1) The $k=0$ vertical transition in germanium, responsible for the "direct" edge in optical absorption at 0.8 eV, takes place between a $\Gamma_{25'}$ valence band maximum and a $\Gamma_{2'}$ conduction band minimum. The transition in the zincblende structure 3–5 compounds which corresponds to $\Gamma_{25'}{}^v \rightarrow \Gamma_{2'}{}^c$ in the diamond structure group 4 elements is $\Gamma_{15}{}^v \rightarrow \Gamma_1{}^c$.[9] The pressure shift of the energy of this $\Gamma_{25'}{}^v \rightarrow \Gamma_{2'}{}^c(\Gamma_{15}{}^v \rightarrow \Gamma_1{}^c)$ transition is within the range $(10-15)\times 10^{-6}$ eV/bar for Ge,[10] GaAs,[11] InSb,[12] and GaSb.[13]

(2) Germanium and many of the 3–5's exhibit doublet peaks in reflectivity in the 2–3 eV range.[14,15] These are due to $\Lambda_3{}^v \rightarrow \Lambda_1{}^c$ vertical transitions along [111] axes in the Brillouin zone.[16] The pressure coefficient of these reflectivity peaks in Ge, GaSb, and InAs is $(7-8)\times 10^{-6}$ eV/bar.[7]

The results for GaP obtained in this study are consistent with this idea of correlation between pressure coefficient and transition assignment.

Comparatively little is known of the band structure of GaP. The valence band maximum is assumed to be a Γ_{15} state at the zone center, in analogy with the other 3–5 compounds.[17–19] The energy gap observed in optical absorption is 2.2 eV at room temperature, 2.33 eV at 0°K.[20–22] Measurements on GaAs–GaP alloys reveal a nonlinear variation of energy gap with composition,[23] indicating that the conduction band minimum in GaP is not the same as in GaAs, i.e., is not a Γ_1 state. This conclusion is supported by the large electron effective mass obtained from mobility and Faraday rotation experiments.[24] That the fundamental absorption edge involves indirect transitions can be seen from the functional form of the dependence of absorption coefficient upon photon energy above threshold,[21] especially the distinct phonon-derived structure observed at low temperatures.[22] The effect of pressure on the band gap is to shift it, slowly, to lower energies,[25] an unusual behavior seen also in silicon.[2,26] This suggests that the conduction band minima, as in silicon, lie along [100] directions in k space. Recent work by Spitzer and Mead on the photoresponse of surface barriers on GaAs-GaP alloys places the direct energy gap at 2.65 eV.[27] Our experiments indicate a somewhat higher estimate for this energy, closer to 2.8 eV, as will be seen in Secs. 3 and 5. Information available from ultraviolet reflectivity measurements and from measurements on an infrared absorption band in n-type material are discussed in later sections, along with our pressure experiments on these properties.

The discussion of our experimental results in Secs. 3–6 is presented in four parts. In Sec. 3 we describe the effect of pressure on the absorption edge between 2.2 and 2.7 eV, showing how the shape of the edge and its pressure dependence implies the presence of indirect transitions of the type $\Gamma_{15}{}^v \rightarrow X_1{}^c$ near 2.2 eV, and direct transitions of the type $\Gamma_{15}{}^v \rightarrow \Gamma_1{}^c$ near 2.8 eV. In Sec. 4 we discuss the 0.3-eV absorption band observed in n-type GaP, its pressure shift, and its interpretation as a direct inter-conduction band transition of the type $X_1{}^c \rightarrow X_3{}^c$. In Sec. 5 we describe the pressure shift of the strong reflectivity peak at 3.7 eV, as well as that of a weak peak at 2.8 eV not hitherto reported. We present evidence for attributing the former to $\Gamma_{15}{}^v \rightarrow \Gamma_{15}{}^c$ and the latter to $\Gamma_{15}{}^v \rightarrow \Gamma_1{}^c$ direct interband transitions. In Sec. 6 we discuss the results obtained for the effect

[9] R. H. Parmenter, Phys. Rev. 100, 573 (1955).
[10] M. Cardona and W. Paul, Phys. Chem. Solids 17, 138 (1960).
[11] J. Feinleib, S. Groves, W. Paul, and R. Zallen, Phys. Rev. 131, 2070 (1963); G. E. Fenner, J. Appl. Phys. 34, 2955 (1963).
[12] D. Long, Phys. Rev. 99, 388 (1955); R. W. Keyes, *ibid.* 99, 490 (1955).
[13] A. L. Edwards and H. G. Drickamer, Phys. Rev. 122, 1149 (1961); T. Deutsch and B. Kosicki, Bull. Am. Phys. Soc. 9, 60 (1964).
[14] H. Ehrenreich, H. R. Philipp, and J. C. Phillips, Phys. Rev. Letters 8, 59 (1962).
[15] H. R. Philipp and H. Ehrenreich, Phys. Rev. 129, 1550 (1963).
[16] D. Brust, J. C. Phillips, and F. Bassani, Phys. Rev. Letters 9, 94 (1962).
[17] F. Bassani and M. Yoshimine, Phys. Rev. 130, 20 (1963).
[18] R. Braunstein and E. O. Kane, Phys. Chem. Solids 23, 1423 (1962).
[19] We shall ignore here the very small terms linear in k near $k=0$, known to be present in the 3–5's, which displace the actual valence band maxima to positions very slightly removed from $k=0$ (see preceding reference).
[20] O. G. Folberth and F. Oswald, Z. Naturforsch. 9a, 1050 (1954); F. Oswald, *ibid.* 10a, 927 (1955); H. Welker, J. Electronics 1, 181 (1955); Doris Teichler (private communication).
[21] W. G. Spitzer, M. Gershenzon, C. J. Frosch, and D. F. Gibbs, Phys. Chem. Solids 11, 339 (1959).
[22] M. Gershenzon, D. G. Thomas, and R. E. Dietz, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 752.
[23] O. G. Folberth, Z. Naturforsch. 10a, 502 (1955).
[24] T. S. Moss, A. K. Walton, and B. Ellis, *Proceedings of the International Conference on Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 295.
[25] A. L. Edwards, T. E. Slykhouse, and H. G. Drickamer, Phys. Chem. Solids 11, 140 (1959); R. Zallen and W. Paul, preliminary work quoted in Ref. 4.
[26] M. I. Nathan and W. Paul, Phys. Rev. 128, 38 (1962); W. Paul and G. L. Pearson, *ibid.* 98, 1755 (1955).
[27] W. G. Spitzer and C. A. Mead, Phys. Rev. 133, A872 (1964).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE II. Optical components used for the various experiments.

Measurement	Spectral region (eV)	Dispersing element	Spectral slit-width (eV)	Source	Detector	Fluid surrounding sample
Fundamental absorption edge	2.2–2.8	glass prism	0.004	W lamp	9526B photomultiplier or 1P21 photomultiplier	isopentane
Infrared absorption band	0.3–0.7	LiF prism	0.003	globar	thermocouple	CS_2
Reflectivity under pressure	2.3–4.5	NaCl prism	0.01	D_2 lamp	9526B photomultiplier	isopentane
Reflectivity at 80°K	1.9–3.2	glass prism	0.004	W lamp	9526B photomultiplier	vacuum
Radiative recombination at 300°K and 80°K	1.7–2.3	glass prism or 7500-Å grating	0.02 0.006	GaP diode	9526B or 7102 photomultiplier	isopentane or He gas

of pressure upon recombination radiation emitted from forward-biased p-n junctions of GaP. The sharp edge emission lines at 2.1–2.3 eV, observed at 80°K, serve as accurate markers for monitoring the shift of the indirect edge. The results obtained in this fashion are compared to the results of Sec. 3.

2. EXPERIMENTAL TECHNIQUES

The spectrometer used in this work was a Perkin-Elmer model 12G unit used as a prism instrument or as a grating instrument in different experiments. Usually, the prism used was glass for the visible region of the spectrum, LiF for the infrared, and NaCl for the ultraviolet. Table II lists the prisms, sources, and detectors used for the various measurements, along with the spectral resolution available in each instance. In some of the low-temperature radiative recombination work, a grating blazed at 7500 Å provided somewhat higher resolution for studying the relatively narrow (0.02-eV half-width) edge emission lines and their small pressure shift. The radiation was chopped mechanically at 13 cps and phase-detected. The resultant ac signal was then amplified, rectified, and recorded on a Leeds & Northrup type G recorder. Data were taken by continuous recording of the signal while scanning photon energy over the region of interest, rather than by a point-by-point method. Optical transmissions down to 10^{-4} were measurable before reaching the limit set by scattered light within the monochromator.

The technique of optical experiments performed at high hydrostatic pressure has been well documented elsewhere.[2,28,29] The pressure fluid was transmitted between the Bridgman press and the sample chamber by means of flexible, stainless-steel tubing.[30] Pressure was measured by means of manganin gauges,[31] calibrated by observing the freezing pressure of mercury

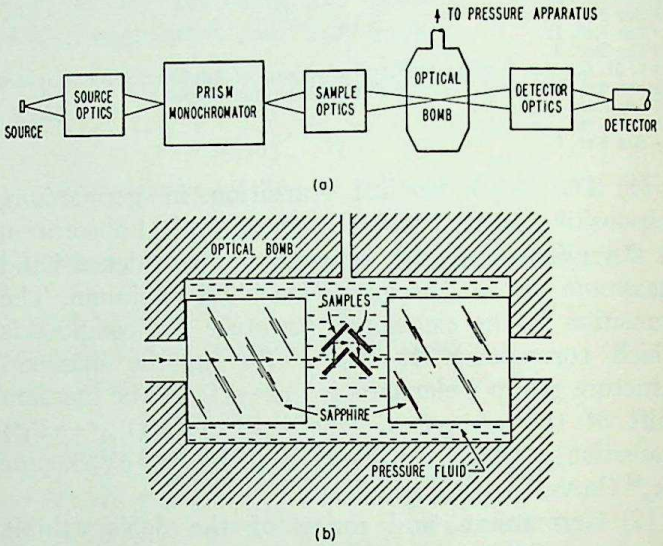


FIG. 1. Experimental arrangement for the measurement of optical absorption or reflectivity as a function of pressure: (a) block diagram of the optical system; (b) schematic diagram of the sample chamber with reflectivity samples in place. For the absorption measurements, the sample was the usual plane-parallel slab.

at 0°C, 7566 bars.[32] Our experiments were conducted at pressures up to 9 kbars.[33]

For the absorption and reflectivity pressure experiments, all performed at room temperature, the pressure fluids used were liquids; isopentane for the visible and ultraviolet measurements, carbon disulfide for the infrared measurements. The sample was placed in an optical high pressure vessel[34] between two $\frac{1}{2}$-in.-thick sapphire windows. A system of plane and spherical mirrors formed a real image of the monochromator exit slit at the sample position. The experimental layout is schematically shown in Fig. 1. For the measurements of absorption coefficient, the sample was a plane-parallel slab with polished surfaces normal to the direction of the light beam. For the measurements of reflectivity, four etched samples were arranged in a periscope-like reflectivity capsule. The transmission of

[28] D. M. Warschauer and W. Paul, Rev. Sci. Instr. 29, 675 (1958).
[29] For a bibliography of 22 references on optical high-pressure techniques, see Ref. 5, p. 442.
[30] W. Paul and D. M. Warschauer, Rev. Sci. Instr. 27, 418 (1956).
[31] P. W. Bridgman, *The Physics of High Pressures* (G. Bell and Sons, London, 1949).
[32] D. K. Newhall, L. H. Abbot, and R. A. Dunn, in *High-Pressure Measurement*, edited by A. A. Giardini and E. C. Lloyd (Butterworth Inc., Washington, 1963).
[33] Throughout this paper, we shall use the bar as our unit of pressure: 1 bar $\equiv 10^6$ dyn/cm^2 = 1.020 kg/cm^2 = 0.987 atm.
[34] D. Langer and D. M. Warschauer, Rev. Sci. Instr. 32, 32 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

this capsule, proportional to R^4, exhibits peaks at photon energies at which the reflectivity peaks occur. Since we were interested in the position in energy of reflectivity peaks, rather than the absolute magnitude of reflectivity, the use of 45° angle of incidence rather than normal incidence does not present any difficulties. This is supported by measurements made with this technique on Ge, InAs, and GaSb,[7] as well as on Si,[8] which yielded values of peak positions in good agreement with values obtained at normal incidence.[35,36] One feature omitted, for reasons of clarity, from Fig. 1, is a set of mirrors which sampled a small fraction of the beam from the monochromator for the purpose of monitoring the stability of the spectral distribution of the energy available in the source-monochromator-detector system.

For the low-temperature measurement of the 2.8-eV reflectivity peak discussed in Sec. 5C, the reflectivity capsule was mounted in the vacuum chamber of a conventional metal cryostat, in thermal contact with a copper block connected to the liquid nitrogen bath.

For the experiments on the emission from GaP diodes, some of which were performed at 80°K, the pressure fluid used was helium gas. This entailed the use of a second pressure system and sample chamber designed for use with gas. The diode was mounted on a combination electrical and optical plug in which a central sapphire window was encircled by electrical lead-in wires. For the low-temperature measurements, the pressure vessel was immersed in liquid nitrogen inside a stainless-steel Dewar. The radiation, emitted through the sapphire window, was conveyed by means of glass fiber optics to the source position of the spectrometer. This arrangement is schematically shown in Fig. 2.

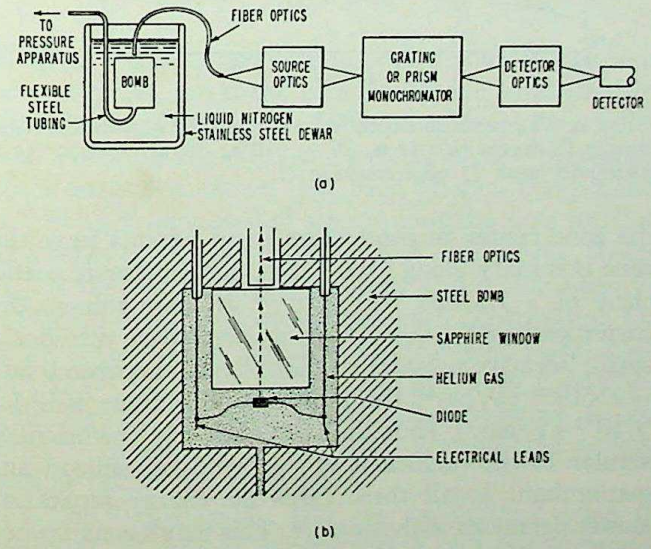


FIG. 2. Experimental arrangement for the measurement of radiative recombination as a function of pressure: (a) block diagram of the optical system; (b) schematic diagram of the sample chamber.

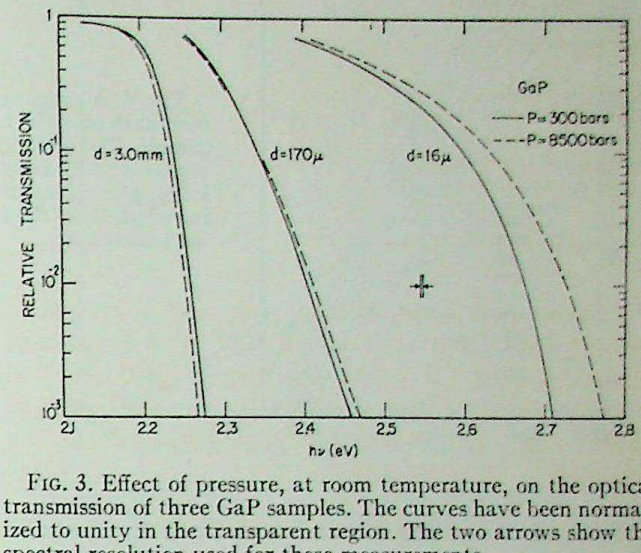


FIG. 3. Effect of pressure, at room temperature, on the optical transmission of three GaP samples. The curves have been normalized to unity in the transparent region. The two arrows show the spectral resolution used for these measurements.

3. EFFECT OF PRESSURE ON THE ABSORPTION EDGE: 2.2–2.7 eV

Transmission samples were prepared by polishing down bulk material. The range of absorption coefficient α covered in our measurements was from 5 to 5000 cm^{-1}. Figure 3 displays representative data taken at room temperature on three samples of different thickness. Relative transmission, t/t_0, where t_0 is the transmission in the transparent region at low energy, is plotted against photon energy for each sample at a low and a high pressure. Pressure shifts the transmission curves in opposite directions for the two extreme samples: There is a small shift toward lower energy for the thick sample, which spans the region of low α for $h\nu$ near 2.2 eV; and a large shift toward higher energy for the thin sample, which spans the region of high α from 2.5 to 2.7 eV. The pressure effect reverses sign within the range covered by the intermediate sample, near 2.33 eV.

Figures 4 and 5 show sets of transmission curves taken during two pressure runs. Figure 4 corresponds to absorption coefficients of the order of 10 cm^{-1}, at photon energies just above the energy gap E_G near the beginning of the absorption edge. Later we will show that the pressure shift of the curves in this region approaches the value of $\partial E_G/\partial P$, the pressure coefficient of the band gap. Figure 5 corresponds to absorption coefficients of the order of 10^3 cm^{-1}, at photon energies well into the fundamental absorption region. The change of shape with pressure is readily apparent; the large positive shift increases rapidly with $h\nu$. The large difference between the ranges of photon energy covered in Figs. 4 and 5 illustrates the difference in magnitude between the small, negative shift with pressure near 2.2 eV, and the large, positive shift with pressure near 2.7 eV.

[35] J. Tauc and A. Abraham, *Proceedings of the International Conference on Semiconductor Physics, Prague, 1960* (Czechslovakian Academy of Sciences, Prague, 1961), p. 375.
[36] M. Cardona, J. Appl. Phys. 32, 2151 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

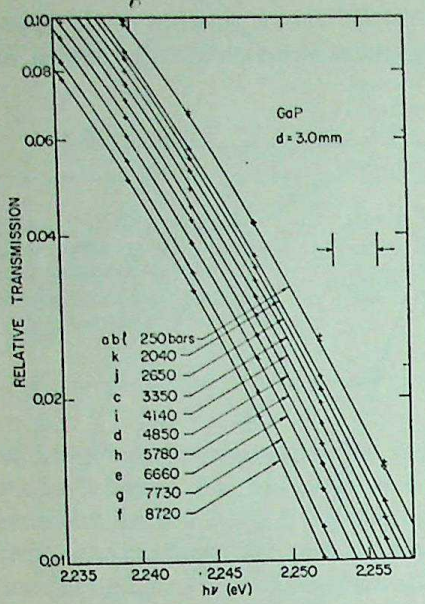


FIG. 4. A pressure run taken on a thick sample. The letters, arranged alphabetically, give the order in which the data were taken.

Plots such as Figs. 4 and 5 were used to construct isoabsorption curves, $h\nu_\alpha = h\nu_\alpha(P)$, in which photon energy for a given relative transmission (i.e., for a given α) is plotted against pressure. These were linear, within our 0–9 kbar pressure range. In Fig. 6 the slopes of these lines, $(\partial h\nu/\partial P)_\alpha$, are shown as a function of their zero-pressure intercepts $h\nu(0)$. The two extreme points, at 2.78 and 2.23 eV, are obtained from other experiments to be discussed in Secs. 5 and 6, respectively. All that need be noted now is that they are consistent with the isoabsorption data spanning the range from 2.24 to 2.72 eV. Figure 6 essentially summarizes the effect of hydrostatic pressure on the room temperature absorption edge of GaP.

To obtain $\partial E_G/\partial P$ from Fig. 6, it is necessary to know the room temperature value of E_G. One method of studying an indirect edge is a high-resolution analysis of phonon-absorption and phonon-emission components of $\alpha(h\nu)$ at low α, as has been successfully performed for Ge and Si.[37] Gershenzon, Thomas, and Dietz[22] have made such a study on GaP at low temperatures, observing phonon-emission components consistent with previous work on phonon combination bands in GaP. Their work is the most direct evidence thus far reported that the energy gap is indirect, and provides the best estimate available for E_G. Although unable to observe phonon structure at room temperature, they extrapolate their low-temperature data to a room-temperature energy gap of 2.22 eV.

Using $E_G = 2.22$ eV, we find from Fig. 6 that

$$\partial E_G/\partial P = -1.2\times10^{-6}\ \text{eV/bar}. \qquad (1)$$

This pressure coefficient is in fairly good agreement with previously reported values, less accurately determined than the present one, although it is somewhat smaller in magnitude.[25] The band gap of silicon, a $\Gamma_{25'}{}^v \rightarrow \Delta_1{}^c$ indirect transition from the valence band maximum at the zone center to conduction band minima near the zone boundary along [100] axes, has a pressure coefficient of -1.5×10^{-6} eV/bar.[2,26] In germanium, the corresponding $\Gamma_{25'}{}^v \rightarrow \Delta_1{}^c$ energy separation also has a small, negative pressure coefficient lying between 0 and -2×10^{-6} eV/bar[3]; the most accurate value is -1.5×10^{-6} eV/bar.[38] Thus the GaP band gap coefficient is similar to the coefficient of $\Gamma_{25'}{}^v \rightarrow \Delta_1{}^c$ in silicon and germanium. In all three cases the energy separation slowly decreases with pressure. This may be contrasted to the behavior of the band gap of germanium, a $\Gamma_{25'}{}^v \rightarrow L_1{}^c$ indirect transition, which increases with pressure at a rate of $+5\times10^{-6}$ eV/bar.[3]

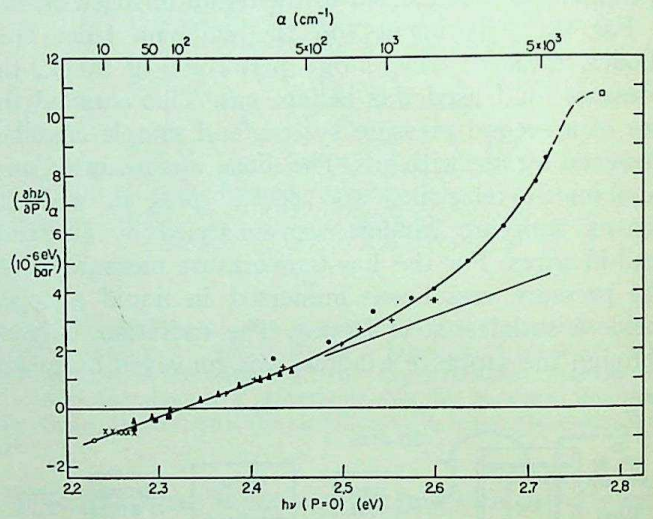


FIG. 6. The pressure dependence of the GaP absorption edge. Sample thickness (μ): 16 ●; 29 +; 170 ▲; 550 ■; 3000 ×. Also, reflectivity peak □, edge emission ○.

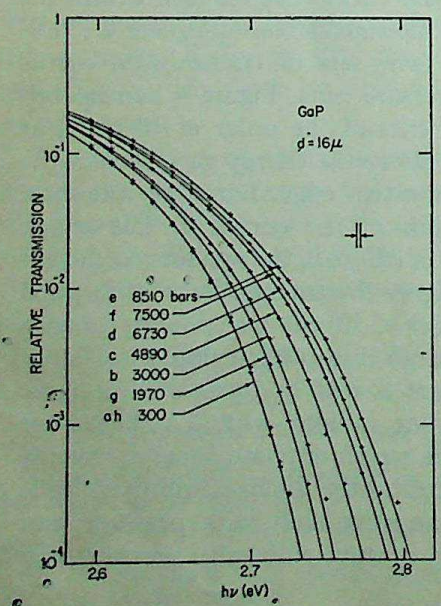


FIG. 5. A pressure run taken on a thin sample. The letters, arranged alphabetically, give the order in which the data were taken.

[37] T. P. McLean, Progr. Semiconductors 5, 53 (1960).
[38] W. E. Howard, Gordon McKay Laboratory of Applied Science, Harvard University, Technical Report HP-7, 1961 (unpublished).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

We conclude then that the indirect transition responsible for the energy gap in GaP is that which corresponds to the $\Gamma_{25'}{}^{v} \rightarrow \Delta_1{}^{c}$ transition in Ge and Si, namely, $\Gamma_{15}{}^{v} \rightarrow \Delta_1{}^{c}$. Since the degeneracy at the X point in the diamond structure is split in the zincblende structure, it is likely that the conduction-band minima in GaP are at the X point on the zone boundary, rather than along the Δ axis in the interior of the zone. Thus we assign the band gap to the transition $\Gamma_{15}{}^{v} \rightarrow X_1{}^{c}$.

Figure 6 clearly shows that the absorption edge does not simply translate rigidly in energy along with E_G when pressure is applied, since for that situation $(\partial h\nu/\partial P)_\alpha$ would be equal to $\partial E_G/\partial P$ for all $h\nu$. Instead, $(\partial h\nu/\partial P)_\alpha$ varies linearly with $h\nu$ between 2.2 and 2.45 eV, beyond which energy it increases progressively more rapidly. Thus, for example, the photon energy corresponding to $\alpha = 5000$ cm^{-1} increases with pressure at a rate of 8×10^{-6} eV/bar, even though E_G itself decreases with pressure.

To account for the variation of $(\partial h\nu/\partial P)_\alpha$ with $h\nu$, we must first discuss the theoretically calculated and the experimentally measured shape of the absorption edge, $\alpha = \alpha(h\nu)$. The simplest theory[37] predicts that for indirect, allowed transitions, α is given by a sum of terms corresponding to phonon emission and absorption processes, and that each term varies quadratically with energy above threshold. Thus,

$$\alpha = \sum_i (\alpha_i{}^{+} + \alpha_i{}^{-}), \tag{2}$$

where:

$$\alpha_i{}^{\pm} = \begin{cases} 0, & \Delta E_i{}^{\pm} < 0 \\ C_i{}^{\pm}(\Delta E_i{}^{\pm})^2, & \Delta E_i{}^{\pm} > 0 \end{cases} \quad \begin{aligned} &\Delta E_i{}^{\pm} = h\nu - E_G{}^{i\pm}, \\ &E_G{}^{i\pm} = E_G \pm k\theta_i. \end{aligned}$$

Here $\alpha_i{}^{+}$ is the component due to emission of a phonon of energy $k\theta_i$, $\alpha_i{}^{-}$ is the component due to absorption of a phonon of energy $k\theta_i$. In the more exact theory,[37,39] which includes the effect of the electron-hole interaction (excitons), it is shown that each phonon component, $\alpha_i{}^{\pm}$, is itself composed of several contributions starting at a series of energies given by $E_G{}^{i\pm} - (1/n^2)E_{ex}$, where E_{ex} is the exciton binding energy and n is a positive integer. For allowed transitions, each contribution for finite n (formation of bound electron-hole pairs) goes as $(\Delta E)^{1/2}$ for small ΔE, the energy above threshold. The highest threshold contribution, starting at $h\nu = E_G{}^{i\pm}$ (formation of unbound electron-hole pairs), goes as $(\Delta E)^{3/2}$ for small ΔE.

When experimental data on an indirect edge is analyzed, it is usual to construct a plot of $\alpha^{1/2}$ versus $h\nu$. From the above it is clear that $\alpha^{1/2}$ versus $h\nu$ would not be expected to reveal simple linear behavior in the complex region near $h\nu \approx E_G$. However, in the region in which $h\nu$ greatly exceeds the highest phonon-emission threshold ($h\nu - E_G{}^{i} \gg E_{ex}$ for all i), the simple theory and the theory including excitons both give the same result:

$$\alpha = C[(h\nu - E_G')^2/h\nu]. \tag{3}$$

The quantity E_G' results from a weighted average of the phonon components and may differ from E_G by an amount of the order of a phonon energy. This expression holds provided that the assumption of parabolic energy bands, used in the theories, is valid to these energies. The factor $(h\nu)^{1/2}$ usually does not vary by more than 10% over the range of interest, so that we might expect $\alpha^{1/2}$ versus $h\nu$ to exhibit a region of linearity at photon energies exceeding $E_G + 0.05$ eV.

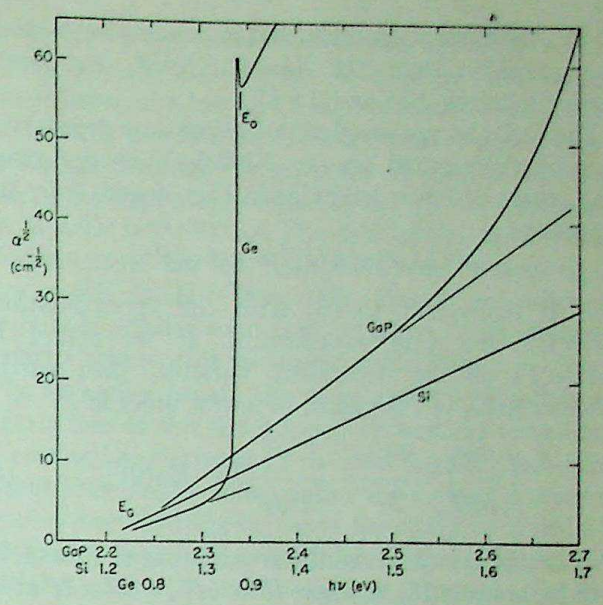


FIG. 7. A comparison of the absorption edges of Ge, Si, and GaP. The curves have been shifted horizontally so that the indirect energy gaps coincide. The Ge data (which are for 20°K) and the Si data (which are for 77°K) are from T. P. McLean, Progr. Semiconductors 5, 53 (1960). The GaP curve is our room-temperature, zero-pressure data.

Our results on the room temperature, atmospheric pressure absorption edge of GaP are in quite good agreement with the data of Spitzer, Gershenzon, Frosch, and Gibbs[21] (henceforth referred to as SGFG). Figure 7 presents a comparison of the absorption edge of GaP with those of Ge and Si. The scales of photon energy have been shifted so that the indirect energy gaps roughly coincide. The vertical scale is the square root of the absorption coefficient $\alpha^{1/2}$. The germanium and silicon data, for 20° and 77°K, respectively, are from the work of Macfarlane *et al.* quoted in Ref. 37. The silicon edge is accurately linear over the range shown, extrapolating to $\alpha^{1/2} = 0$ at a photon energy of 1.19 eV, 0.03 eV higher than the actual band gap at 77°K. The germanium edge is linear over a small range of energy until the onset of direct transitions at E_0, the direct energy gap. The extrapolated energy is 0.74 eV, in agreement with E_G (to the nearest 0.01 eV). The GaP edge is linear below about 2.5 eV, extrapolating to zero absorption at 2.20 eV, 0.02 eV lower than the band gap as estimated by Gershenzon *et al.* Above

[39] R. J. Elliott, Phys. Rev. 108, 1384 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

2.5 eV, the absorption coefficient increases progressively more rapidly than the $(h\nu-2.2\text{ eV})^2$ dependence obeyed between 2.25 and 2.5 eV.

The GaP absorption edge and its pressure dependence may be represented, in the 2.2–2.5-eV energy range (the region of linear behavior in Figs. 6 and 7) by the expression

$$\alpha^{1/2}(h\nu,P)=A(P)[h\nu-E_G(P)]. \quad (4)$$

This is equivalent to (3) with the approximations $E_G'=E_G$ and $(h\nu)^{-1/2}=$constant. If we regard $h\nu=h\nu(\alpha,P)$ as the dependent variable, then partial differentiation of (4) yields, to lowest order in P,

$$\left(\frac{\partial h\nu}{\partial P}\right)_\alpha=\frac{\partial E_G}{\partial P}-\frac{1}{A}\frac{\partial A}{\partial P}[h\nu-E_G(0)]. \quad (5)$$

This is just the observed linear variation of $(\partial h\nu/\partial P)_\alpha$ with $h\nu$ shown in Fig. 6, where $(\partial h\nu/\partial P)_\alpha=\partial E_G/\partial P$ at $h\nu=E_G(0)$, and where the slope of the linear region, $(d/dh\nu)(\partial h\nu/\partial P)_\alpha\equiv\beta$, is given by $-(1/A)(\partial A/\partial P)$. In Sec. 7, the quantity β is estimated theoretically from the proposed pressure-dependent band structure of GaP. The result is in satisfactory agreement with the measured value of 1.15×10^{-5} bar^{-1} obtained from Fig. 6.

In the region of photon energy above 2.5 eV, the curves of Figs. 6 and 7 deviate from their linear behavior at lower energy. Both $\alpha^{1/2}$ and $(\partial h\nu/\partial P)_\alpha$ increase, gradually at first and then more and more rapidly, above the line extrapolated from low energies. The suggestion which immediately comes to mind is that a new mechanism, strongly pressure-dependent, is beginning to contribute to the absorption above 2.5 eV.

Several possible explanations for the rise in absorption are listed below:

(a) the breakdown of the validity of the theory behind (3), due to such considerations as the k dependence of matrix elements, energy denominators, and especially, effective masses (nonparabolic energy bands away from the extrema);

(b) the onset of a new indirect transition process at 2.5 eV, involving, for example, a set of higher conduction band minima;

(c) the low-energy "tail" of a direct edge occurring at a higher energy not reached in our absorption measurements;

(d) the onset of direct transitions at 2.5 eV.

Both (a) and (b) are plausible possibilities. The valence and conduction band masses may increase at k vectors away from the extrema, thereby providing a joint density of states for indirect transitions exceeding that predicted on the basis of parabolic bands. Also, regarding (b), it will be seen in Sec. 4 that there is evidence for a higher conduction band lying, at the X point, about 0.3 eV above the lowest conduction band. However, our calculations indicate that these two mechanisms are probably too weak to account for the observed increase in α above 2.5 eV. Note, for example, that (a) is apparently unimportant in the case of silicon, for which $\alpha^{1/2}$ versus $h\nu$ is linear over the 0.5-eV range of photon energy shown in Fig. 7. If the additional contribution to the absorption of GaP is analyzed by constructing $\alpha_2\equiv\alpha-\alpha_1$, where α_1 is extrapolated from the low-energy region where $\alpha=\alpha_1=A(h\nu-2.2\text{ eV})^2$, then it is found that α_2 does not at all resemble the $(h\nu-2.5\text{ eV})^2$ dependence expected of an indirect transition beginning at 2.5 eV, but increases much more rapidly than this. Thus, neither (a) nor (b) accounts for the observed shape of the edge. Moreover, neither mechanism readily accounts for the more rapid increase of $(\partial h\nu/\partial P)_\alpha$ with $h\nu$ above 2.5 eV.

If the minimum "vertical" separation between the valence and conduction bands in GaP occurs at $k=0$ and the transition involved is of the type $\Gamma_{15}^v\rightarrow\Gamma_1^c$, as is the case in several other 3–5 compounds, then the pressure coefficient of the direct gap, $\partial E_0/\partial P$, is expected to have a value in the range of $+10$ to $+15\times10^{-6}$ eV/bar. Mechanisms (c) and (d) would then predict an increase of $(\partial h\nu/\partial P)_\alpha$ toward the value of $\partial E_0/\partial P$. In germanium it is found that $(\partial h\nu/\partial P)_\alpha$ reaches the value of the direct gap coefficient at an energy somewhat less than E_0, and then levels off.[10] For GaP we see that $(\partial h\nu/\partial P)_\alpha$ is still below the expected value of $\partial E_0/\partial P$ at $h\nu=2.7$ eV, indicating that E_0 occurs at a higher energy. Thus, only (c) is consistent with the variation of $(\partial h\nu/\partial P)_\alpha$ of Fig. 6. SGFG proposed (d) to account for the increase in α above 2.5 eV, and to account for an infrared absorption band observed in n-type material. Our discussion in this section favors (c) as the explanation for the increase in α; the actual origin of the infrared absorption band will be discussed in Sec. 4.

In a direct absorption edge, the absorption coefficient rises rapidly to a high value at E_0, and then increases more slowly with $h\nu$.[37,40] The direct edge of Ge is shown in Fig. 7, and clearly exhibits the characteristic "knee" at E_0. This curve is for 20°K and shows the exciton peak at an energy just under E_0. At room temperature the exciton peak is washed out but the marked knee at E_0 persists. No indication of such a knee appears in the GaP absorption edge of Fig. 7. This also demonstrates that E_0 does not fall within the range shown, and that it thus exceeds 2.7 eV.

From these arguments we therefore propose that (c) is the dominant mechanism responsible for the deviation from linearity above 2.5 eV in Figs. 6 and 7. The presence of a $\Gamma_{15}^v\rightarrow\Gamma_1^c$ direct transition is indicated at an energy E_0 greater than 2.7 eV. The momentum matrix element for this transition, p_{cv}, which governs the optical absorption, is nearly equal for several 3–5 compounds.[41] Assuming that the same value holds also

[40] M. D. Sturge, Phys. Rev. 127, 768 (1962).
[41] H. Ehrenreich, J. Appl. Phys. 32, 2155 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

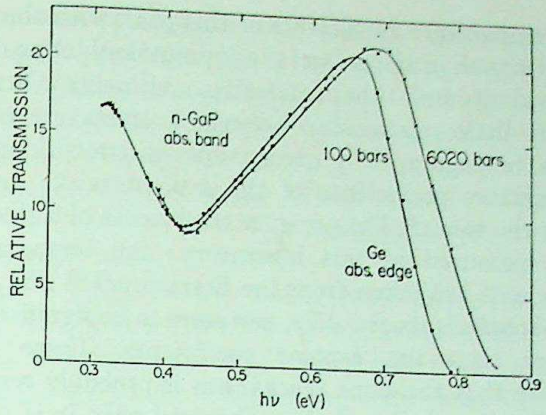


FIG. 8. Effect of pressure on the infrared absorption band in n-type GaP. The trough at 0.4 eV is due to the GaP absorption band. The cutoff at 0.7–0.8 eV is due to the absorption edge of a Ge sample which was included along with the GaP sample in order to provide a comparison of pressure shifts.

for GaP, we can obtain a rough estimate for $\alpha(E_0)$, the absorption level at the knee of the absorption curve. This quantity is proportional to $p_{cv}^2 m_c^2/n^3 E_0$,[37,39] where n is the refractive index, and where it is assumed thta both the density-of-states effective mass and the exciton reduced mass are determined primarily by m_c, the conduction band effective mass. In first approximation, m_c is proportional to E_0,[41] so that $\alpha(E_0)$ becomes proportional to $p_{cv}^2 E_0/n^3$. For GaAs, $\alpha(E_0)$ is equal to 0.9×10^4 cm^{-1}.[40] Using the values of n and E_0 for GaAs and GaP, where a trial value of $E_0=3$ eV is assumed for GaP, we obtain $\alpha(E_0)\approx2\times10^4$ cm^{-1} for GaP. The absorption curve of Fig. 7 extrapolates to this level at a photon energy of about 2.75–2.80 eV. In Sec. 5C we shall discuss a weak peak appearing in the reflectivity spectrum of GaP at about 2.78 eV, which we attribute to the $\Gamma_{15}^v \rightarrow \Gamma_1^c$ direct edge.

4. EFFECT OF PRESSURE ON AN INFRARED ABSORPTION BAND IN n-TYPE GaP

SGFG reported an absorption band in n-type GaP, in excess of the normal free carrier absorption, with a strength proportional to the electron concentration.[21] At room temperature, the band peaks at 0.41 eV and its low-energy threshold is at 0.31 eV; at low temperatures the threshold shifts to higher energy by about 0.05 eV. Since the absorption is proportional to the concentration of electrons, the initial states of the transitions responsible are probably either at the conduction band minima, or at shallow impurity states derived from these minima.

Spitzer *et al.* suggested that the mechanism responsible is an indirect transition from the conduction band minima to a higher minimum at $k=0$, i.e., $X_1^c \rightarrow \Gamma_1^c$. As a test of this hypothesis, which is inconsistent with our previously discussed estimate of the energy of Γ_1^c, we measured the effect of pressure on the absorption band in samples kindly provided us by Dr. Spitzer. Since $\partial E_0/\partial P$ exceeds $+10\times10^{-6}$ eV/bar, and $\partial E_G/\partial P$ is negative, the pressure shift of E_0-E_G (the energy of the $X_1^c \rightarrow \Gamma_1^c$ transition) should definitely exceed $+10\times10^{-6}$ eV/bar. Figure 8 displays the optical transmission, at a low and a high pressure, of an n-type GaP sample ($n=10^{18}$ cm^{-3}, $d=0.1$ mm) in series with a Ge sample included for purposes of comparison. The shift of the high-energy cutoff near 0.75 eV, due to the Ge, is $+9\times10^{-6}$ eV/bar. The shift of the dip in transmission caused by the absorption band is seen to be very much smaller than this, almost an order of magnitude smaller than that expected. This is strong evidence against this particular mechanism.

In an earlier paper, one of us (W.P.) presented an explanation of this band based on vertical transitions of conduction electrons to the next higher conduction band.[4] This simple model, unlike the indirect scattering process discussed above, leads naturally to a definite low-energy threshold, especially if the electrons are initially weakly bound to shallow impurities. It also explains reasonably well the magnitude of the absorption band and its relative insensitivity to temperature, as no phonons are involved. The model is consistent with a small pressure coefficient. Recently, Allen and Hodby[42] have studied this absorption band in n-type GaAs-GaP alloys. They found that it appears only in the alloys which have the (100) minima lowest, strong evidence in favor of our model.

The higher conduction band providing the final states for these direct transitions is logically X_3^c, which is degenerate with X_1^c in diamond-structure materials like Ge and Si. Thus the $X_1^c \rightarrow X_3^c$ energy separation in GaP is about 0.3 eV. Values for this quantity for some of the other 3–5's have been estimated by Greenaway[43] from structure observed at low temperatures in the strong X-point reflectivity peak in the ultraviolet. He obtains values in the range of 0.4–0.5 eV, close to the GaP value. The interconduction band absorption in n-type GaP which has been discussed here is the analog of the intervalence band infrared absorption found in p-type Ge,[44] GaAs,[45] and InAs;[46] the shape of the absorption curve is similar in all of these materials.

5. THE EFFECT OF PRESSURE ON THE REFLECTIVITY SPECTRUM: 2.5–4.5 eV

A. Reflectivity Spectra

Reflectivity spectra of semiconductors exhibit peaks in the 1–10-eV range of photon energies, due to corners (peaks or edges) in dN/dE, the joint density-of-states for direct interband transitions.[14–16] Corners in dN/dE are correlated with critical points in k space of $E^{cv}(k)$

[42] J. W. Allen and J. W. Hodby, Proc. Phys. Soc. 82, 315 (1963).
[43] D. L. Greenaway, Phys. Rev. Letters 9, 97 (1962).
[44] H. B. Briggs and R. C. Fletcher, Phys. Rev. 87, 1130 (1952); *ibid.* 91, 1342 (1953).
[45] R. Braunstein, Phys. Chem. Solids 8, 280 (1959).
[46] F. Stern and R. M. Talley, Phys. Rev. 108, 158 (1957).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$=E^c(k)-E^v(k)$, the vertical separation of the conduction and valence bands. The critical points are of four types; minima, maxima, and two types of saddle points. The saddle points produce the strongest singularities in dN/dE, and it is their identifications with particular reflectivity peaks which are the most certain. The strongest reflectivity peak in Ge, Si, and the 3–5 compounds appears at about 4 or 5 eV and is due to saddle point singularities at X and Σ points[47] in the zone.[14,16] Doublet peaks appearing in the 2–3 eV range in Ge and some of the 3–5's are due to saddle points at Λ points[47] in the zone,[16] the doublet structure being due to the spin-orbit splitting of the valence band.

Reflectivity peaks have been reported for GaP at 3.7, 5.3, 6.9, and 10 eV.[15] The dominant 5.3-eV peak is supposed due to the X, Σ critical point transition.[14,16] The $L_3^v \rightarrow L_3^c$ (or $\Lambda_3^v \rightarrow \Lambda_3^c$) transition has been suggested for the 6.9-eV peak. No transition has been associated with the peak at 10 eV. The 3.7-eV peak may be due to $\Lambda_3^v \rightarrow \Lambda_1^c$ transitions, in analogy with the 2–3-eV peaks in Ge and some of the 3–5's; or it may be due to $\Gamma_{15}^v \rightarrow \Gamma_{15}^c$ transitions, in analogy with the $\Gamma_{25'} \rightarrow \Gamma_{15}^c$ transitions believed responsible for the first peak in Si at 3.4 eV.[14,36]

In this section we shall discuss measurements which we have made on the reflectivity between 2.5 and 4.5 eV. We have measured the effect of pressure on the peak at 3.7 eV, as well as on a weak peak at 2.8 eV which has not been reported previously.

B. The Reflectivity Peak at 3.7 eV

Figure 9 shows the effect of pressure on the 3.7-eV peak. The vertical scale is the transmission of a reflectivity capsule (four successive 45° reflections from etched surfaces) of GaP immersed in isopentane, and is proportional to R^4. The peak position shifts toward higher energies at the rate of $+(5.8\pm15\%)\times10^{-6}$ eV/bar. In Table III we have compared the pressure and temperature coefficients of this peak with those of the first peak in silicon and the Λ-point doublet peaks in germanium and other materials. Although there is usually little relationship between temperature coefficients belonging to a given type of transition, the temperature coefficients of the Λ-point peaks are all about the same.[48] The pressure coefficients of Table III were measured in this laboratory; the temperature coefficients are taken from the literature.[36,48] The GaP and Si coefficients are alike, and seem to be significantly smaller than the Λ-point coefficients. These data indicate that the same mechanism is probably responsible for the GaP and Si peaks, and that it is not a Λ-point transition.[49]

That the Si 3.4-eV peak is other than a Λ-point transition has already been conclusively established by work of Tauc and Abraham[50] on Ge-Si alloys, in which it was shown that the variation of the energy of the first reflectivity peak (followed as a function of composition) shows a definite break at about 79% silicon. It now remains to identify this peak, and thereby the GaP 3.7-eV peak. There is a burgeoning body of evidence that the $\Gamma_{25'}^v \rightarrow \Gamma_{15}^c$ energy separation in silicon is

TABLE III. Comparison of the behavior of the 3.7-eV reflectivity peak in GaP with that of some other reflectivity peaks.

Coefficient	$\Lambda_3 \rightarrow \Lambda_1$ peaks at 2–3 eV in Ge, GaAs, GaSb, etc.	Peak at 3.7 eV in GaP	Peak at 3.4 eV in Si
$(\partial E/\partial P)_T$ in 10^{-6} eV/bar (±10%)	+7.2[a]	+5.8	+5.2[b]
$(\partial E/\partial T)_P$ in 10^{-4} eV/°K (±20%)	−4.7[c,d]	−2.5[c]	−2.6[c,d]

[a] See Ref. 7. [b] See Ref. 8. [c] See Ref. 36. [d] See Ref. 48.

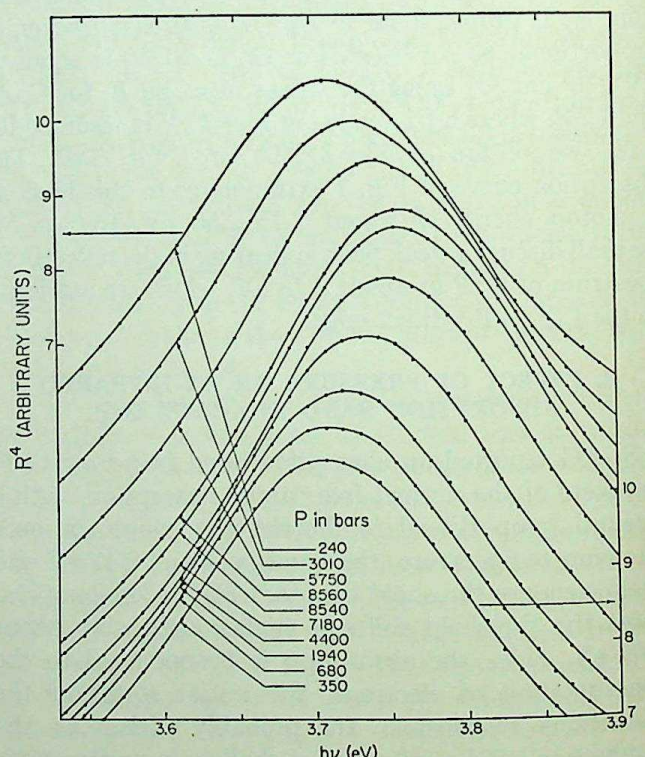


FIG. 9. Effect of pressure on the GaP reflectivity peak at 3.7 eV. The curves have been shifted vertically in order to separate them. The pressures are listed in the order in which the data were taken.

[47] We are using the accepted notation for various positions of high symmetry in the Brillouin zone of a face-centered cubic lattice: Γ, (000); Δ, along [100] axes; X, (100) at the zone boundary; Σ, along [110] axes; K, (110) at the zone boundary; Λ, along [111] axes; L, (111) at the zone boundary.

[48] F. Lukes and E. Schmidt, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 389.

[49] It is worth noting that the lack of observed doublet structure in the Si and GaP peaks does not, by itself, make the Λ-point explanation unlikely, since the expected spin-orbit splitting in these materials is small, less than 0.1 eV (Ref. 18).

[50] J. Tauc and A. Abraham, Phys. Chem. Solids 20, 190 (1961).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

3.4 eV.[51–54] The band calculations for silicon seem to have successfully accounted for the reflectivity spectrum,[52] photoemission experiments on silicon surfaces,[52,53] and the alloy experiments mentioned above. We therefore assign the 3.4-eV Si peak to $\Gamma_{25'}{}^{v} \rightarrow \Gamma_{15}{}^{c}$. We conclude then that the 3.71-eV reflectivity peak in GaP is due to $\Gamma_{15}{}^{v} \rightarrow \Gamma_{15}{}^{c}$ transitions ($\Gamma_{25'}{}^{v}$ in Si corresponds to $\Gamma_{15}{}^{v}$ in GaP).

GaP belongs to the isoelectronic sequence GeSi, GaP, ZnS, CuCl. All of these materials exhibit very similar reflectivity spectra,[15,50,55,56] each peak occurring at an energy which is higher the more polar the compound. The data of Tauc and Abraham[50] indicate that the energies of the $\Gamma_{25'}{}^{v} \rightarrow \Gamma_{15}{}^{c}$ and $\Lambda_3{}^{v} \rightarrow \Lambda_1{}^{c}$ transitions in GeSi are 3.1 and 2.9 eV, respectively. Cardona,[56] from work on CuBr-CuCl alloys, puts the corresponding energies in CuCl at 6.8 and 6.5 eV. In both materials only a single reflectivity peak is observed, the weaker peak apparently being hidden under the stronger (which is less than 0.3 eV away). Cardona suggests that this is likely also the case in the other two members of the series. This would place the $\Lambda_3{}^{v} \rightarrow \Lambda_1{}^{c}$ energy separation in GaP in the 3.5–4.0-eV range where, in reflectivity, it is hidden under the (in this case) stronger $\Gamma_{15}{}^{v} \rightarrow \Gamma_{15}{}^{c}$ peak at 3.7 eV. A study of reflectivity in the alloy system GaAs-GaP would determine whether or not this is the correct picture. Finally, corroborating evidence for the Γ-point assignment to the 3.7-eV peak in GaP comes from the corresponding peak in ZnS which shows no splitting under the use of polarized light, and is therefore due to transitions at $k=0$.[57]

C. The Reflectivity Peak at 2.8 eV

In Sec. 3 we estimated a value of about 2.8 eV for E_0, the $\Gamma_{15}{}^{v} \rightarrow \Gamma_1{}^{c}$ direct gap. This type of direct edge, at which there is a sharp rise (edge) in dN/dE due to a minimum in $E^{cv}(k)$, is much weaker than other direct transitions: Typical values of α for this edge are of the order of 10^4 cm^{-1},[37,40] while typical values of α in the region of (for example) the $\Lambda_3{}^{v} \rightarrow \Lambda_1{}^{c}$ and the $X_4 \rightarrow X_1$ transitions are of the order of 10^6 cm^{-1}.[15,58] It is therefore not surprising that for Ge, no structure has yet been reported in the reflectivity spectrum near $E_0=0.8$ eV. For GaAs, there is some evidence of weak structure near $E_0=1.5$ eV.[46] Structure near 2.8 eV has not been present in published reflectivity spectra of GaP.

We have examined the transmission of a reflectivity capsule of GaP in the vicinity of 2.8 eV. Figure 10 shows the transmission (proportional to R^4) of the capsule in vacuum at 300 and 80°K. [It was found that the arrangement of the samples accidentally placed a small transmission sample (thickness about 1 mm) effectively in parallel with the R^4 samples. This is responsible for the small additional contribution to the transmission below about 2.3 eV. The shift of the cutoff of this added component yields a value of $-(5.2\pm15\%)\times10^{-4}$ eV/°K for $\partial E_G/\partial T$, the temperature coefficient of the indirect gap, in agreement with a previously reported value[20] of -5.4×10^{-4} eV/°K.] Superimposed upon the rapid rise of R^4 toward the 3.7-eV peak, there is very weak structure near 2.8 eV. This takes the form of a peak having a height of the order of $\Delta R^4/R^4\approx1/30$, and a width at half-maximum of about 0.2 eV. The temperature coefficient of the peak position is $-(4.6\pm15\%)\times10^{-4}$ eV/°K.

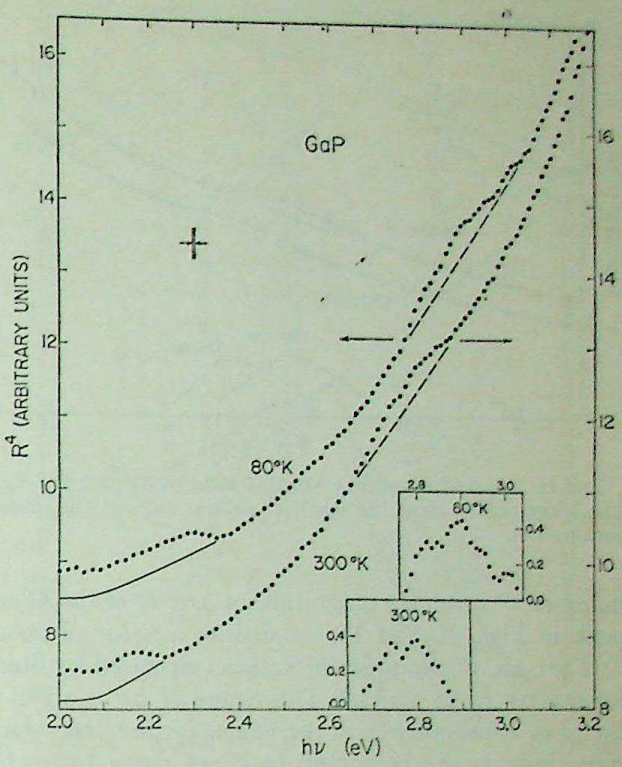


FIG. 10. Effect of temperature on the GaP reflectivity peak at 2.8 eV. The inserts show the data obtained by subtracting out the smooth background.

Figure 11 shows the effect of pressure on the 2.8-eV "peak," measured at room temperature, with the samples immersed in isopentane. Its pressure coefficient is $+(10.7\pm15\%)\times10^{-6}$ eV/bar. This large coefficient is appropriate for a $\Gamma_{15}{}^{v} \rightarrow \Gamma_1{}^{c}$ direct transition.[3–6] It has been plotted as the high-energy point in Fig. 6 (at 2.78 eV, the zero-pressure photon energy of the peak), where it is seen to be consistent with the absorption edge data obtained at lower energies.

Structure in reflectivity in this region (where the extinction coefficient, $K\equiv(C/2\omega)\alpha\approx10^{-5}$ cm$\times\alpha$, is less than unity) closely follows structure in the refractive

[51] J. C. Phillips, Phys. Rev. 125, 1931 (1962).
[52] D. Brust, M. L. Cohen, and J. C. Phillips, Phys. Rev. Letters 9, 389 (1962).
[53] W. E. Spicer and R. E. Simon, Phys. Rev. Letters 9, 385 (1962).
[54] F. Bassani and D. Brust, Phys. Rev. 131, 1524 (1963).
[55] W. C. Walker and J. Ostantowski, J. Opt. Soc. Am. 53, 399 (1963); M. Cardona, Phys. Rev. 129, 1068 (1963).
[56] M. Cardona, Phys. Rev. 129, 69 (1963).
[57] M. Cardona, Solid State Communications 1, 109 (1963).
[58] H. R. Philipp and E. A. Taft, Phys. Rev. 113, 1002 (1958); 120, 37 (1960).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

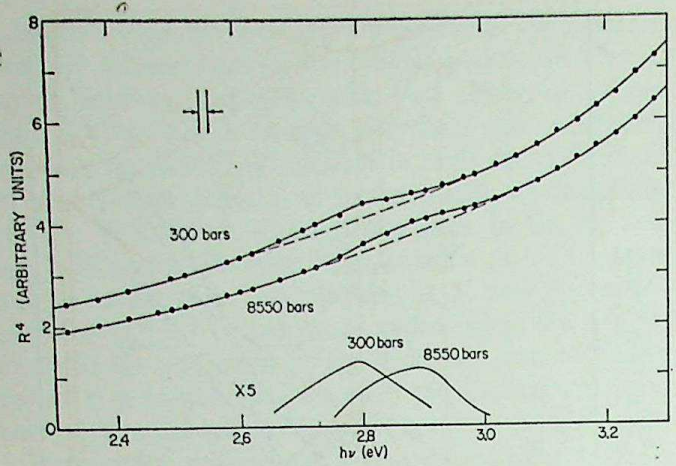


FIG. 11. Effect of pressure on the GaP reflectivity peak at 2.8 eV. The lower curves show the result of subtracting out the smooth background.

index n.[58,59] From the magnitude of $\Delta R^4/R^4$ of the 2.8-eV peak in Figs. 10 and 11, we deduce a value of about 0.02 for Δn, of the order of values calculated by Stern for n near E_0 in GaAs.[60] This value of Δn is about a factor of 10^2 smaller than the values for the first strong reflectivity peaks belonging to edges rising to $\alpha \approx 10^6$ cm^{-1},[58] implying that the edge involved has a strength of about 10^4 cm^{-1}.

From this evidence we conclude that this weak reflectivity peak is due to the onset of direct transitions of the type $\Gamma_{15}^v \rightarrow \Gamma_1^c$ at the center of the zone. The pressure and temperature dependence of the direct gap is then given by

$$E_0 = 2.78 + (10.7 \pm 15\%) \times 10^{-6} P - (4.6 \pm 15\%) \times 10^{-4} (T - 300), \quad (6)$$

where E_0 is in electron volts, P is in bars, T is in degrees Kelvin. Recently, Spitzer and Mead have estimated a value for E_0 of 2.65 eV based on measurements of the photoresponse of metal-semiconductor surface barriers on GaAs-GaP alloys.[27] We do not know the reason for the difference between our two results.

The photoconductivity of GaP shows a sharp drop at about 2.8 eV.[61] Similar effects have been observed in other materials,[62] and have been accounted for by the production of photoelectrons at $k=0$, where they contribute little to the current because of rapid direct recombination.[63] This is consistent with our estimate of E_0.

Note added in proof. D. F. Nelson, L. F. Johnson, and M. Gershenzon [Bull. Am. Phys. Soc. 9, 236 (1964)] have recently observed structure in the photoconductivity spectra of GaP samples which they attribute to direct exciton formation. Their room-temperature estimate of E_0, 2.79 eV, is in excellent agreement with our result, as is the temperature dependence which they obtain for E_0. The agreement between these two independent determinations of the direct gap is strong evidence for the correctness of our interpretation of this reflectivity peak, as discussed above.

6. EFFECT OF PRESSURE ON RADIATIVE RECOMBINATION IN FORWARD-BIASED p–n JUNCTIONS

Diodes formed from GaP p-n junctions exhibit carrier-injection electroluminescence (EL) emission of low efficiency when biased in the forward direction.[64,65] At low temperatures the emission in the green, at or near E_G, reveals a great deal of very complicated and interesting fine structure[65–67] due to such effects as associated shallow donor-acceptor pairs[66] and bound exciton states.[67] There is also a broad emission in the red involving levels well inside the forbidden zone.[64]

We have measured the effect of pressure on some of the EL lines observed in GaP diodes. A reason for doing this is that sharp lines emitted with nearly the band-gap energy serve as accurate and convenient markers for monitoring the pressure shift of the band gap, as has been demonstrated for GaAs.[11] We can thus obtain in straightforward fashion an independent measurement of $\partial E_G/\partial P$ to compare with the result of the analysis given in Sec. 3 of the pressure dependence of $\alpha(h\nu)$.

Table IV summarizes our data on two green-emitting diodes fabricated at Bell laboratories by diffusing Zn into nonintentionally doped, epitaxially grown, n-type material. The diodes were provided to us by Dr. M.

TABLE IV. Effect of pressure on GaP radiative recombination lines emitted at near band-gap energy.

Temp. (°K)	Emission line[a,b]	$h\nu$ (eV)	$\Delta h\nu$[c] (eV)	Relative intensity[d]	$-\partial h\nu/\partial P$ (10^{-6} eV/bar)
80	EE–2LO	2.096	0.04	6	...
80	EE–LO	2.145	0.04	25	1.0±0.1
80	EE	2.194	0.04	100	1.0±0.1
80	A–Ac–LO	2.238	0.02	60	1.2±0.1
80	A–Ac	2.294	0.02	10	1.4±0.5
80	A	2.308	0.02	2	1±1
300	...	2.23	0.1	...	1.1±0.1

[a] See Ref. 65.
[b] EE is edge emission, A is an exciton line, LO is longitudinal-optical phonon, Ac is acoustical phonon.
[c] Approximate full width at half-maximum.
[d] At a current of 10 mA.

[59] B. Velicky, Czech. J. Phys. B11, 787 (1961).
[60] F. Stern, Phys. Rev. 133, A1653 (1964).
[61] L. F. Johnson and D. F. Nelson, unpublished data reported in Ref. 22; H. G. Grimmeiss, W. Kischio, and H. Koelmans, Solid-State Electron. 5, 155 (1962).
[62] E. Loh, Phys. Chem. Solids 24, 493 (1963).
[63] E. Loh and J. C. Phillips, Phys. Chem. Solids 24, 495 (1963).
[64] M. Gershenzon and R. M. Mikulyak, Solid State Electronics 5, 313 (1962); J. Starkiewicz and J. W. Allen, Phys. Chem. Solids 23, 881 (1962).
[65] M. Gershenzon, R. M. Mikulyak, R. A. Logan, and P. W. Foy, Solid-State Electron. 7, 113 (1964).
[66] J. J. Hopfield, D. G. Thomas, and M. Gershenzon, Phys. Rev. Letters 10, 162 (1963); D. G. Thomas, M. Gershenzon, and F. A. Trumbore, Phys. Rev. 133, A269 (1964).
[67] D. G. Thomas, M. Gershenzon, and J. J. Hopfield, Phys. Rev. 131, 2397 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Gershenzon, who, along with co-workers, has carefully studied the EL spectra of such units and compared them with photoluminescence spectra.[65] The labeling of the various lines given in Table IV follows the discussion of Gershenzon, Mikulyak, Logan, and Foy,[65] and was obtained by comparison of our spectra with the similar 78°K spectra of their Fig. 8.

The notation for the 80°K lines of Table IV is as follows: EE signifies edge emission and is ascribed to recombination at widely separated donor-acceptor pairs.[65,66] It is associated with the sharp-line pair-spectra which appear at higher energies. The A line has been ascribed to annihilation of an exciton bound to an impurity.[65,67] The other four low-temperature lines are replications of the EE and A emissions at lower energies due to the simultaneous emission of phonons, acoustic (Ac), or longitudinal optical (LO). The edge emission lines saturate at higher currents while the A lines do not. The room temperature emission is either band-to-band or involves very shallow states.

Although the results for the weakest lines listed in Table IV are relatively less certain than the others, all yield about the same pressure coefficient. The consensus value for $\partial h\nu/\partial P$ is $-(1.1\pm0.1)\times10^{-6}$ eV/bar, in agreement with the band gap coefficient obtained in Sec. 3 from the absorption edge data. The room temperature datum of Table IV provided the lowest energy point shown in Fig. 6. We found no appreciable effect of pressure on the intensity of any of the lines.

The validity of equating the pressure coefficients of these emission lines with $\partial E_G/\partial P$ depends only on the assumption that the transitions involved are between shallow states derived from the band edges, which seems reasonable since their energies lie so close to E_G. However, it has been suggested that the A lines are due to a direct exciton (derived from the valence and conduction band states at $k=0$) strongly bound to an impurity, and that they accidentally occur at an energy near E_G because of a large binding energy. If this were so, then their pressure coefficient should be close to $\partial E_0/\partial P = +11\times10^{-6}$ eV/bar, rather than to $\partial E_G/\partial P = -1\times10^{-6}$ eV/bar. Table IV shows that the reverse is true. Thus, provided that we have correctly identified these lines as belonging to the A complex of Refs. 65 and 67, the mechanism for the A lines more probably involves an ordinary indirect exciton.

We have also measured, at room temperature, the effect of pressure on the broad emission band (width about 0.2 eV) appearing at 1.8 eV. This red emission has been attributed to radiative transitions from a deep donor level, associated with the presence of oxygen in the crystal,[56] to the valence band. The pressure coefficient is $+(1.0\pm0.5)\times10^{-6}$ eV/bar. This coefficient is slightly surprising; it is opposite in sign to $\partial E_G/\partial P$. [In silicon it was found that the pressure coefficient of the energy separation between a deep level and either band edge was a fraction of the (negative) band gap coefficient.[26]] This result implies that the donor binding energy decreases more rapidly with pressure than the band gap.

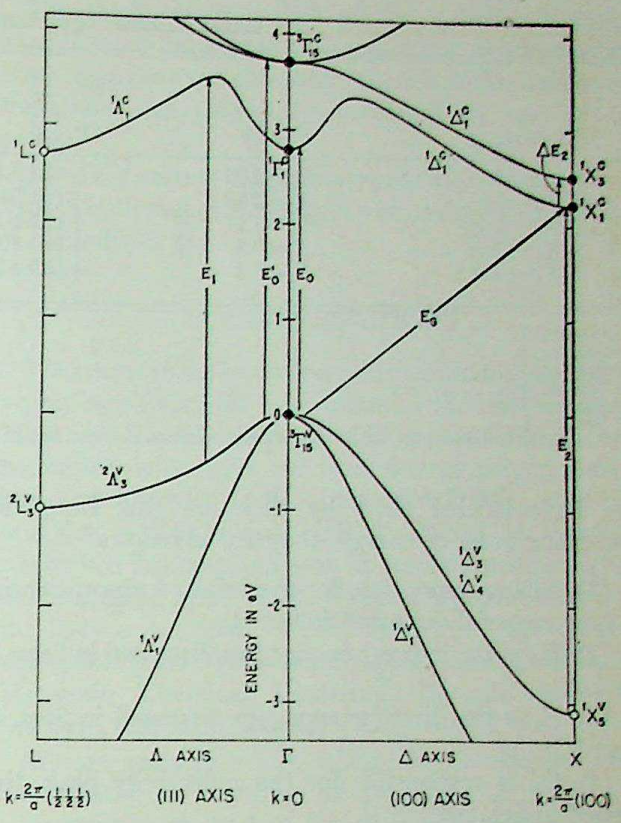


FIG. 12. Band structure of GaP. The states are labeled using the notation for the irreducible representations of the single group of the zincblende lattice. The left-hand superscript appended to each symbol gives the orbital degeneracy of the state. The arrows represent transitions observed in optical experiments. The five energy levels marked by solid circles were those studied by means of our pressure experiments.

7. THE BAND STRUCTURE AND ITS PRESSURE DEPENDENCE

In Fig. 12 we have constructed an electronic energy band diagram for GaP based upon available experimental information. The five energy levels marked by the solid circles are those upon which our pressure experiments bear. The three marked by open circles are from other data discussed below, the L-point levels being relatively uncertain. The levels at the symmetry points have been connected by plausible smooth curves which, although somewhat arbitrary, are consistent with various band structure calculations on similar semiconductors, as well as such considerations as crossing or noncrossing of bands and relative magnitudes of effective masses. We have neglected the spin-orbit splitting of valence band levels, which for GaP is about 0.1 eV. Since we do not consider effects of the electron spin, we have labeled the states with the irreducible representations for the single group, rather than for the double group, of the zincblende structure.[6]

The arrows shown in the figure represent the transitions which have been observed in optical experiments.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

TABLE V. Effect of pressure on the band structure of GaP.

Transition	E in eV	$\frac{\partial E}{\partial P}$ in 10^{-6} eV/bar	Measurement
$\Gamma_{15}^v \rightarrow X_1^c$	$E_G=2.22$	$-(1.1\pm10\%)$	absorption edge,[a] recombination radiation[b]
$\Gamma_{15}^v \rightarrow \Gamma_1^c$	$E_0=2.78$	$+(10.7\pm10\%)$	absorption edge,[c] reflectivity peak[d]
$\Gamma_{15}^v \rightarrow \Gamma_{15}^c$	$E_0'=3.71$	$+(5.8\pm10\%)$	reflectivity peak[e]
$X_1^c \rightarrow X_3^c$	$\Delta E_2=0.3$	$+(1\pm50\%)$	absorption band[f]
$X_5^v \rightarrow X_1^c$	$E_2=5.3$	$+(2.9\pm20\%)$[g]	...

[a] See Sec. 3.
[b] See Sec. 6.
[c] See Sec. 3.
[d] See Sec. 5C.
[e] See Sec. 5B.
[f] See Sec. 4.
[g] This is the coefficient for the corresponding transition is silicon, as measured by means of the 4.3 eV silicon reflectivity peak (see Ref. 8).

We have labeled them by their energies. Listed in ascending order of energy, they are as follows:

(1) ΔE_2 is responsible for the infrared absorption in n-type material discussed in Sec. 4;

(2) E_G is the indirect energy gap discussed in Secs. 3 and 6;

(3) E_0 is the direct energy gap discussed in Secs. 3 and 5C;

(4) E_0' is responsible for the reflectivity peak discussed in Sec. 5B;

(5) E_1 is the Λ-point transition, hidden under E_0', discussed in Sec. 5B;

(6) E_2 is responsible for the 5.3-eV reflectivity peak mentioned in Sec. 5A.

In this work, we have studied the effect of pressure on (1) through (4).

The values of E_G, E_0, E_0', E_2, and ΔE_2, are sufficient to fix the energies (relative to ${}^3\Gamma_{15}^v$, the valence band maximum) of ${}^1X_1^c$, ${}^1\Gamma_1^c$, ${}^3\Gamma_{15}^c$, ${}^2X_5^v$, and ${}^1X_3^c$, respectively. To estimate the energies of ${}^1L_1^c$ and ${}^2L_3^v$, we made use of the following assumptions:

(1) E_1 is equal to E_0' (E_0' is the stronger transition which prevents E_1 from being seen);

(2) the separation of the two bands at L is not very different from E_1 (true for Ge[16]);

(3) ${}^2L_3^v$ is about 1 eV lower than ${}^3\Gamma_{15}^v$. (The known values of E_0, E_1, and the relative energy of the Γ and X conduction band minima, put the $\Gamma-L$ valence band width at 1.1 eV in GaSb, 1.5 eV in Ge.[7])

The two L-point energies shown in Fig. 12 are then specified. In view of the assumptions needed for their computation, they are clearly much less certain than the other levels.

The right-hand side of Fig. 12 is very similar to a band structure which has been calculated for AlP.[17] AlP is the 3–5 compound which is isoelectronic with Si. This points up the fact that GaP more closely resembles Si than it does Ge. Calculations on Ge-Si alloys[68] show that it is the silicon-rich alloys which most closely correspond to the GaP band structure of Fig. 12.

[68] F. Bassani and D. Brust, Phys. Rev. 131, 1524 (1963).

Our experiments determined the effect of pressure on E_G (Secs. 3, 6), E_0 (Secs. 3, 5C), E_0' (Sec. 5B), and ΔE_2 (Sec. 4). The pressure coefficients are listed in Table V. We have also listed a value for $\partial E_2/\partial P$. This is the pressure coefficient, as measured by us, for the corresponding peak in Si (at 4.3 eV).[8]

In the discussion of Sec. 3 on the pressure dependence of the indirect edge, we defined a quantity $(d/dh\nu)(\partial h\nu/\partial P)_\alpha=\beta$. The measured value of β, obtained from the slope of the linear, low-energy region of Fig. 6, was 1.15×10^{-5} bar^{-1}. We can now compare this experimental result with a theoretical estimate of β based, under certain assumptions, upon the information given in Table V.

In Sec. 3, we showed that β is equal to $-(1/A)\times(\partial A/\partial P)$, where A is the proportionality factor of Eq. (4). In the second-order perturbation theory used in dealing with indirect transitions, it is shown that the oscillator strength for this process [given by the coefficients $C_i^{\pm}$ of Eq. (2), C of Eq. (3), and A^2 of Eq. (4)] is proportional to a term of the form

$$\left(\sum_j \frac{p_j}{E_j}\right)^2;$$

p_j is a matrix element and E_j is an energy denominator corresponding to a virtual two-step process involving a particular intermediate state, specified by j.[69] The pressure dependence of this term is dominated by the pressure dependence of the energy denominators. If we assume, for lack of better information, that all of the p_j's are equal, it follows that $A=\text{const}\cdot\sum_j E_j^{-1}$. Substituting this into the expression for β, we find that

$$\beta=\frac{\sum_j(1/E_j^2)(\partial E_j/\partial P)}{\sum_j 1/E_j}. \tag{7}$$

[69] We have assumed that phonon energies are negligible relative to the energy separations of electron states ($k\theta_i \ll E_j$ for all i, j). Otherwise E_j^i would depend on a phonon energy (specified by i), as well as on the energy of the intermediate electronic state (specified by j). The expression for $C_i^{\pm}$ would then just contain E_j^i in place of E_j, but the situation for C and A^2 would be more complicated.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In Fig. 12 there are four intermediate states available for the indirect transition labeled by E_G: ${}^1\Gamma_1{}^c$, ${}^3\Gamma_{15}{}^c$, ${}^1X_3{}^c$, and ${}^2X_5{}^v$. The corresponding energy denominators are E_0-E_G, $E_0'-E_G$, $E_G+\Delta E_2$, and E_2-E_G, respectively. Using (7) and the information given in Table V (and ignoring all other intermediate states but the ones lying within the energy range of Fig. 12), we obtain $\beta=1.0\times10^{-5}$ bar^{-1}. The excellent agreement with the experimental value of 1.15×10^{-5} bar^{-1} is most likely fortuitous in view of the drastic assumptions made; nevertheless, it does show that the measured β is consistent with the proposed pressure-dependent band structure of Fig. 12 and Table V.

8. SUMMARY

The effect of pressure on several optical properties of GaP has been used to construct a tentative electronic energy-band structure for this material over a region of energy within about 4 eV of the valence band maximum at $k=0$. The major experimental results are summarized in Table V; the deduced band structure is shown in Fig. 12.

The data of Table V are consistent with the idea of correlation between pressure coefficients and types of electronic transitions[4-6]: The coefficients of E_G and E_0' are close to those for the corresponding transitions in Si[2,26]; that of E_0 is close to the corresponding coefficient in Ge[10] and GaAs.[11] The result for E_G confirms the location of the lowest conduction band minima as lying along [100] directions in k space.

Evidence for the direct gap E_0 comes from absorption measurements at lower energies, as well as from the weak reflectivity peak at about 2.78 eV, which has not been previously reported. Its temperature coefficient is about -4.6×10^{-4} eV/°K, compared to about -5.2×10^{-4} eV/°K for the indirect gap.

The small pressure shift of the 0.3-eV absorption band observed in n-type material indicates that vertical transitions between conduction bands is the mechanism responsible, as previously proposed by one of the authors.

All of the narrow edge emission lines we observed from a green-emitting diode shifted with pressure at the rate of the indirect gap.

ACKNOWLEDGMENTS

We are indebted to several laboratories for generously supplying us with the GaP material used in this study. Most of the absorption edge measurements were performed on samples cut from a large polycrystalline ingot provided by J. W. Allen of S.E.R.L., who also provided us with two red-emitting diodes. M. Gershenzon of Bell Laboratories sent us the two green-emitting diodes studied at 80°K. W. G. Spitzer provided us with the n-type material which exhibited the infrared absorption band. Other material was received from the Monsanto Chemical Company, Battelle Memorial Institute, and RCA Research Laboratory.

We are grateful to Dr. Allen, Dr. Gershenzon, and Dr. Spitzer for sending us preprints of their own interesting work, and to Dr. M. Cardona of RCA for valuable discussions. We have also benefited from discussions with P. Grant and V. Prakash of this laboratory.

The optical samples used for absorption and reflectivity measurements were prepared by David MacLeod. The optical high-pressure vessel was constructed largely by the efforts of James Inglis. The help of Albert Manning in the construction of apparatus is also acknowledged. We thank B. Kosicki for assisting with some of the diode experiments.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Thermal Disorder in a Dislocation Core and its Effect on Damping*

P. D. SOUTHGATE AND K. S. MENDELSON
Physics Division, IIT Research Institute, Chicago, Illinois

(Received 27 January 1964)

It is postulated that, at elevated temperatures, the core of a dislocation consists of a thermally disordered quasiliquid region, whose size increases with temperature. Calculations based on the effect of the disordering in relieving elastic stresses indicate that a flat strip core should be favored. A moving dislocation of this form will experience two distinct damping forces. One is thermodynamic, caused by the heat of disordering, and is analogous to the thermoelastic damping produced by the dislocation stress field; while the other is viscous, caused by the shear between the two sides of the strip. The calculated magnitudes of both of these damping forces in covalent and metallic crystals are comparable with those required to explain internal friction observations, although a positive demonstration of the validity of the mechanisms cannot be made on the basis of present data.

INTRODUCTION

A DISLOCATION moving through a crystal experiences a damping force, the magnitude of which depends on the dislocation velocity. Measurements of this damping force must be made indirectly and interpretation of the experiments is often open to doubt. It is generally supposed, however, that for small velocities, corresponding to values of the applied shear stress that are small compared to the shear modulus, the damping force is proportional to the velocity. This is the region which will be of interest in this paper.

Theoretically it is found that the nature of the damping force depends very strongly on the type of model assumed for the dislocation. This dependence is particularly sensitive to the structure of the dislocation core, where the Hooke's law approximation breaks down. The simplest model considers the core to be of very small radius and to move as an approximately straight cylinder. If the remainder of the crystal is treated as a continuum exhibiting linear elastic behavior, the only forces are those acting through the long-range stress field of the dislocation. Thermoelastic damping was treated by Eshelby[1] using this model in conjunction with a classical macroscopic thermodynamic method of calculation. If, in addition, nonlinear elastic effects occur near the core, interaction with thermal vibrations will occur. The periodic nature of the crystal can be taken into account by considering the dislocation core to be affected by a periodic Peierls stress field. Leibfried[2] and Nabarro[3] considered such a model, assuming that acoustic waves were scattered by a straight dislocation moving through the lattice without change of form.

A further improvement in the model is made by considering that the dislocation core does not move in the Peierls field as a straight segment, but rather proceeds by the formation and lateral motion of kinks.[4] The main sources of damping are then the energy losses experienced in thermal activation of kinks and the damping of the individual kinks. At low temperatures the damping is caused by the diffusion of built-in kinks (kinks necessitated by the geometry of the dislocation[5]); as the temperature is raised a little the thermal activation of kinks becomes important.[6–10] It appears logical to use discrete kink models at these moderately low temperatures. As the temperature becomes higher, however, the kink density becomes large and kink interactions become frequent. Eventually so much interaction occurs that one can no longer identify individual kinks. In this situation it is better to treat the entire region of the dislocation occupied by kinks as a disordered region of the crystal, and to consider this region as the dislocation core.

Such a model of the dislocation is discussed in this paper. It is assumed that the disorder in the dislocation core is very similar to that found in the liquid state, while the crystal outside the core is treated as a solid isotropic elastic continuum. There are then two separate sources of energy loss which originate at the core of the moving dislocation; one thermodynamic, due to the heat of disordering of the material passing through the dislocation core, and the other viscous, due to the shear rate of the disordered material. In the next section we calculate the size of the disordered region. Following sections then contain calculations of the magnitudes of the two types of damping, and a discussion of the applicability of the theory to experiment.

EQUILIBRIUM SIZE OF DISORDERED CORE

An immediate indication that disordering must occur in the core of a dislocation near the melting point is

* This work supported by Aeronautical Systems Division, U. S. Air Force.

[1] J. D. Eshelby, Proc. Roy. Soc. (London) A197, 396 (1949).
[2] G. Leibfried, Z. Physik 127, 344 (1950).
[3] F. R. N. Nabarro, Proc. Roy. Soc. (London) A209, 278 (1951).
[4] W. T. Read, Jr., *Dislocations in Crystals* (McGraw-Hill Book Company, Inc., New York, 1953), p. 46.
[5] A. D. Brailsford, Phys. Rev. 122, 778 (1961).
[6] J. Lothe and J. P. Hirth, Phys. Rev. 115, 543 (1959).
[7] J. Lothe, Phys. Rev. 117, 704 (1960).
[8] A. Seeger and P. Schiller, Acta Met. 10, 348 (1962).
[9] G. Donth, Z. Physik 149, 111 (1957).
[10] V. Celli, M. Kabler, T. Ninomiya, and R. Thomson, Phys. Rev. 131, 58 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

obtained by considering the lowering of the melting point of the material as a consequence of hydrostatic pressure produced by the stress field of an edge dislocation. In a purely elastic and isotropic material, the pressure within an off-axis cylinder of diameter r will exceed $Gb/2\pi r(1-\nu)$, where G is the shear modulus, b the Burgers vector, and ν Poisson's ratio. The sign depends on whether the cylinder lies above or below the glide plane. If one now applies the Clapeyron equation for melting of bulk material, the melting point will be depressed in the positive pressure region for a material that contracts on melting, or in the tension region if the material expands. For example, in silicon, a material which has geometrically well-defined dislocations at low temperatures, a region of diameter of the order 10 Å will have its melting point suppressed more than 10% (170°C). In this paper we will refer to the region which becomes melted or disordered as the "core." In reality the core will not retain a cylindrical shape, since the material there will not be able to support shear stresses, and hence redistribution of the general stress pattern will occur, favoring a different core shape.

The exact method of calculating the core size would be to consider the free energy of the whole dislocation and then minimize this. Such a calculation would be very elaborate, involving a continuously decreasing disorder at increasing distances from the center. A more limited model is considered here, in which the core is taken to be a single phase, uniformly disordered, immediately surrounded by ideally elastic material. The core size may now be derived in a manner analogous to that employed for the standard Clapeyron equation.

Consider the core to have fixed, but arbitrary, shape and variable size. Denote the core size by a parameter ρ, having dimensions of length. The independent thermodynamic variables are then ρ and the temperature T. Initially it is supposed that the core is acted upon by an arbitrary stress field, whose net effect will be represented by a parameter p, conjugate to ρ, which may be thought as of analogous to a pressure which produces melting point depression. Take a boundary immediately surrounding the core, of dimension ρ_0, and let the internal energy of the total material within the boundary be $U(\rho,T)$. Then the heat entering through the boundary, in any general change, will be

$$dQ=dU+dW=(\partial U/\partial T)dT+(\partial U/\partial\rho)d\rho+pd\rho. \quad (1)$$

If these changes take place reversibly, the entropy change dQ/T will be a perfect differential. Hence,

$$(\partial U/\partial\rho)+p=T(\partial p/\partial T). \quad (2)$$

A change $d\rho$ involves the disordering of a volume $f(\rho)d\rho$ of material, with latent heat of solidification L. In an isothermal change, this heat all passes through the boundary. A change of temperature will require transfer of specific heat also; so, in general,

$$dQ=cdT+Lf(\rho)d\rho. \quad (3)$$

Comparing (1) and (3) and noting that $c=\partial U/\partial T$, yields

$$f(\rho)L=(\partial U/\partial\rho)+p \\ =T(\partial p/\partial T) \text{ from (3)}. \quad (4)$$

In fact the material transforming is in a state of strain in the ordered phase, and in addition the total interface area may be changed during transformation. The effective latent heat is hence that for unstrained material, L_0, less the strain energy and surface energy contributions $L_s(\rho)$.

At this point we restrict consideration to equilibrium conditions only, with the fixed dislocation stress field as the source of p. Now ρ and p are functions of T only, and the partial differential in (4) may be replaced by a total. Integrating, and inserting the sum in place of L, gives

$$p/(L_0-L_s)f(\rho)=\ln T+\text{const}. \quad (5)$$

If now the elastic strain energy outside the boundary is $U_0(\rho,T)$, the parameter p is given by

$$p=\partial U_0/\partial\rho,$$

since all changes of elastic energy appear as work done on the core.

The constant in (5) may be evaluated by considering the dislocation at the melting point. Here the core remains melted in the absence of any peripheral forces, and hence the elastic system should be in a state of static equilibrium. The stress-field energy will continuously reduce as the core size increases, while the surface energy of the solid-liquid interface will increase. The latter is also of the nature of an elastic energy, and so the equilibrium core size ρ_m at T_m may be obtained by minimizing the total stress-field plus surface energy.

Two possible configurations of the disordered core, representing extreme cases, have been considered. They are the flat strip, lying in the glide plane (or some minimum energy plane such as the most closely packed, for a screw dislocation), and the cylinder. It is shown in the Appendix that the rate of change of elastic energy outside a fixed contour close to the core, as the core size changes, is finite for the strip, and zero for the cylindrical configuration. It is probable that the strip represents the most efficient configuration for the relief of stress energy by the core disordering mechanism, and it will be assumed from now on that the core takes this form.

Let q be the thickness and 2ρ the width of the strip, and suppose that only ρ varies with temperature. This constraint will affect the final answer only slightly, and is introduced as a simplifying approximation. The total elastic and surface energy (see Appendix) is

$$(G\beta b^2/4\pi)\ln(kR/\rho)+4\rho s,$$

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where s is the energy per unit area of liquid-solid interface, and the constant β varies from $1/(1-\nu)$ for an edge dislocation to unity for a screw. Minimizing this gives

$$\rho_m = Gb^2\beta/16\pi s.$$

Noting that the shape function $f(\rho)$ will be $2q$, and substituting in Eq. (5) yields

$$\rho = \frac{Gb^2\beta}{8\pi([L_0-L_s]q\ln(T_m/T)+2s)} \tag{6}$$

for the equilibrium core size in the flat strip configuration.

The approximate stress distribution discussed in the Appendix gives a shear strain in the solidifying material at the strip edges of $b\beta/2\pi q$, and hence the strain energy modification to L is

$$Gb^2\beta^2/8\pi^2\rho^2.$$

This turns out usually to be only a small fraction of L_0, and so it may be introduced into (6) as a second-order correction. The magnitude of the surface energy correction is not certain, but it is assumed that this also is fairly small.

DAMPING FORCE

We now calculate the magnitudes of the thermodynamic and viscous damping forces mentioned in the Introduction, again assuming that in most respects the situation may be treated as macroscopic. Both sources of damping are additional to any which may act on the dislocation stress field.

The thermodynamic damping is derived from the rate of entropy gain due to the temperature difference between the leading and trailing edges of the core. If the core were sufficiently wide, the temperature difference would be proportional to the normal thermal resistivity as well as the rate of latent heat release. However, the width is usually comparable with the mean free path of the thermal phonons, and in this situation the heat-transfer mechanism becomes one of phonon radiation rather than diffusion. The phonons may be considered to travel with an effective group velocity v_s, rather less than that of lower frequency acoustic waves. If the rate of heat generation per unit length of dislocation at the solidifying edge is Qq, the excess heat density near the source will be Q/v_s. The temperature difference between leading and trailing edges is therefore

$$\delta T = 2Q/av_sc_v,$$

where c_v is the specific heat per unit volume, and a is an effective radius, of the order of one atom distance, of the material which may be considered to have its temperature changed.

The entropy gain rate, treating the system macroscopically, is

$$\frac{dS}{dt} = Q\left(\frac{1}{T_1}-\frac{1}{T_2}\right) = \frac{Q\delta T}{T^2},$$

if δT is small. If the configuration of the dislocation core remains constant as it moves, the rate of change of internal energy of the system will be

$$dU/dt = T(dS/dt) = Q\delta T/T, \tag{7}$$

which may be represented by a damping force coefficient B_t:

$$dU/dt = B_tV^2, \tag{8}$$

where V is the dislocation velocity. The heat-generation rate will be

$$Q = (L_0-L_s)qV, \tag{9}$$

where L_0-L_s is the effective latent heat of solidfication. Thus, from (7), (8), and (9), the damping coefficient becomes

$$B_t = 2(L_0-L_s)^2q^2/aTv_sc_v. \tag{10}$$

Viscous damping may be calculated very simply for a strip-shaped core by assuming that the shear flow pattern is laminar. The relative displacement of the two faces as the core passes by is b; since it takes time $2\rho/V$ to pass, the shear rate is

$$\dot{\theta} = (b/q)(V/2\rho).$$

The total power dissipation is $2\rho q\eta\dot{\theta}^2 = B_vV^2$, giving

$$B_v = \eta b^2/2\rho q. \tag{11}$$

In the case of a core which has a shape closer to the cylindrical, the flow pattern might become more complex. There will be a tendency for the core to "roll" as the dislocation progresses, lessening the frictional force. However, estimates of the flow pattern have been made for a cylinder, allowing the formation of four main eddies, which show that the viscous damping differs from Eq. (11) by a factor of the order of unity. It is supposed that this effect will be unimportant in the case of a nearly flat strip.

APPLICATION OF MODEL

The theory developed in the previous sections should apply to dislocations in crystals which have a low-impurity content so that interactions between impurities and dislocations are not significant, and in which dislocation motion is over such short distances that they do not actually intersect other dislocations in the course of their motion. Internal friction measurements, or microscopic studies of the motion of identifiable dislocations, are two ways in which such a situation may be studied. Unfortunately there is a scarcity of experimental data, the tendency having been to con-

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

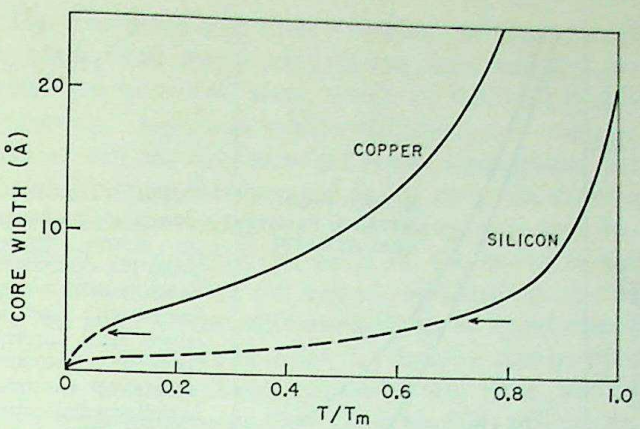


FIG. 1. Widths of disordered cores computed for copper and silicon. Arrows denote Burgers vector of dislocation, below which size the width loses its significance.

centrate on the behavior of dislocations at lower temperatures where appreciable dislocation broadening is not expected. We will take the case of two materials, for which there is some published information: silicon and copper. These form fairly typical covalent and metallic crystals. Ionic crystals are expected to behave in a similar manner to metallic, since the parameters involved have similar values.

The most significant factor in the theory is the heat of disorder. As a first approximation, this could be equated to the heat of fusion of the material and in copper we will assume this is the value. In silicon, melting involves not only disordering but also a change from essentially covalent to essentially metallic bonding,[11] which change presumably is exothermic. If we assume that in the core the disorder is sufficiently localized so that the covalent bonding is retained, and that one bond per atom is broken on the average, requiring about 1.5 eV, a heat of disorder is derived which is approximately twice that of melting; this will be used.

The second factor involved is the surface energy of the interface, and the value of this is very uncertain. No direct measurements exist, and such values as are quoted are derived from observations of supercooling of solidifying melts, using nucleation theory.[12] It is not clear whether the surface energy of significance under these conditions is the same as that at the dislocation core. We shall therefore guess a value for the surface energy of one-tenth the disorder energy, per atom. The core width will be sensitive to error in this approximation only near the melting point.

TABLE I. Values used in computing core widths.

	Silicon	Copper
Latent heat L (ergs cm^{-3})	1.2×10^{11}	1.6×10^{10}
Shear modulus G (dyne cm^{-2})	8×10^{11}	7.5×10^{11}
Burgers vector b (Å)	4.0	2.55
Strip width q (Å)	6.0	4.0
Melting point (°K)	1685	1356

Figure 1 shows the variation of core width to be expected for silicon and copper as a function of temperature. The dislocations considered are simple 60° dislocations in the (111) glide plane in both cases; Table I lists the values of parameters used. An isotropic average shear modulus is taken. It can be seen that as the temperature reduces to the Bordoni peak region in copper, below 110°K, and the region below 1200°K in silicon, the dislocation becomes less than one Burgers vector wide and it may be expected that the Peierls force should attain its full value. Above these temperatures, a smearing-out effect will occur. In both materials, dislocation behavior below these temperatures has been explained in terms of kink mechanisms, which require a large Peierls force; hence, the disordered-core theory is consistent with the kink hypothesis.

The thermodynamic component of damping may be computed from Eq. (10), the assumption being made that the effective phonon velocity is of the order 10^5 cm/sec. In silicon, B_t is then 2.5 dyn-sec-cm^{-2} at 1000°K and in copper 10^{-2} dyn-sec-cm^{-2}. These values are considerably larger than those obtained from the other models in which the Peierls force is not important, such as the thermoelastic theory of Eshelby ($B\sim10^{-5}$) or the interaction of thermal vibrations via nonlinear behavior near the core ($B\sim10^{-4}$), and hence will always dominate these other mechanisms.

Viscous damping is more difficult to evaluate, since a considerable extrapolation of viscosity data to the supercooled case is required. We assume that the viscosity is a thermally activated process, with the coefficient of viscosity $\eta\sim10^{-4}\exp(H/kT)$. In copper, the activation energy is assumed similar to that in other metals,[13] about 0.1 eV. In silicon, a higher activation energy is taken, corresponding to a bond-breaking process; the value of 1.4 eV gives a good match to the experimental observations. The viscous damping exceeds the thermodynamic damping below 0.1 T_m in copper and 0.85 T_m in silicon. It appears, therefore, that in copper (and, presumably many other metals also) the viscous mechanism is never important, since below 0.1 T_m the kink description of dislocation motion becomes valid. In silicon, on the other hand, it is anticipated that it forms a good description between 0.7 T_m and 0.85 T_m.

At the present time, comparison with experiment cannot do more than indicate support for the theory as applied to silicon. A major difficulty is the extraction of a damping coefficient from the measurements of

[11] D. R. Hamilton and R. L. Seidensticker, J. Appl. Phys. 34, 2697 (1963).

[12] J. H. Holloman and D. Turnbull, in *Progress in Metal Physics*, edited by B. Chalmers (Interscience Publishers, Inc., New York, 1953), Vol. 4, p. 333.

[13] C. J. Smithells, *Metals Reference Book* (Butterworths Scientific Publications, Ltd., Washington, 1962), Vol. II, p. 698.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

internal friction, since the free length of dislocation is an important, and frequently unknown, factor. The Koehler-Granato-Lucke[14,15] stretched-string model of dislocation damping gives for the logarithmic decrement δ of a crystal, at low oscillation amplitudes,

$$\delta=\frac{\alpha 8Eb^2NL^2}{\pi^3 C}\frac{\omega\tau}{1+\omega^2\tau^2},$$

with relaxation time

$$\tau=L^2B/\pi^2C,$$

where E is Young's modulus, N the dislocation density, L the free length of the dislocation, C its line tension, and α an orientation factor of the order 0.25. Now if $\omega\tau\gg1$, both L and C vanish from the expression, and B can be computed readily from δ and N. For silicon, where the disordered-core theory predicts B always greater than unity, the inequality will hold for measurements at 10 kc/sec as long as $L>4\times10^{-4}$ cm, which will be the case in pure material with dislocation densities of less than 10^7 cm^{-2}. Figure 2 gives a curve for B derived from measurements of δ on a silicon crystal bar oscillating at 8 kc/sec and containing a dislocation density of approximately 10^6 cm^{-2}. The form of the curve is very similar to that obtained by summing B_t and B_v from the theory; the factor of 2 or so difference in magnitude is not significant as it could easily arise from a nonuniform orientation distribution of a portion of the dislocation.

In the case of copper, the situation is more complex as the damping coefficient is expected to be lower, and hence the dislocation length will become important. For most measurements reported, it is probable that $\omega\tau<1$. The internal friction will then be proportional to the damping coefficient, which itself varies inversely as the temperature. Such behavior is not observed, increases with temperature always being reported above about 200°K, sometimes gradual,[16] sometimes rapid.[17] Two additional effects, which would influence the temperature dependence of the logarithmic decrement, might be mentioned here. The first is the possible release of dislocation pinning points as the temperature increases. This could cause an increase of decrement with temperature, either by increasing the free dislocation lengths, or by allowing some less limited motion at the pinning points. Secondly, dislocations in the (111) plane of face-centered cubic metals split into two half-dislocations. Changes with temperature in the spacing of these half-dislocations, particularly the possibility of their recombining at higher temperatures,

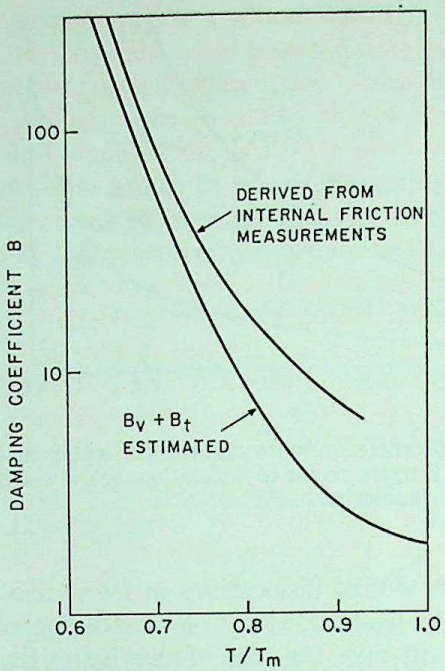


FIG. 2. Dislocation damping coefficient in silicon derived from the disordered-core model, compared with that derived from internal friction measurements by assuming a uniform dislocation distribution and free length greater than 10^{-3} cm.

could affect the temperature dependence of the logarithmic decrement. It is not clear what sort of effect this might be, although it is not expected to modify the behavior drastically. It is evident that further theoretical and experimental investigation is required.

DISCUSSION

A thermally disordered dislocation is a complicated system whose behavior can only be specified accurately in terms of the motions of a large number of atoms. The situation lies midway between a macroscopic and microscopic system; too small for the averaging processes implicit in the treatment of macroscopic systems to be used with complete rigor, and yet so large that analysis in terms of microscopic variables such as local phonon modes, or the statistics of some disorder parameter such as the kink density, would become immensely complex. The model which has been presented here uses the macroscopic approximation, modified at the points where it obviously breaks down, such as in the thermal conductivity. However, the validity of this approximamation is purely intuitive, since the treatment is, strictly speaking, inconsistent; neither can it be said at present to be supported by any good fit to experimental data. In its defense, it may be argued that the dislocation is at least of macroscopic length, so that although thermal fluctuations, for instance, may exceed the systematic difference due to dislocation motion over a very short length of line, over the total length they will be averaged out.

One prediction of the theory is that even in an isotropic material, a screw dislocation will lose its radial symmetry at high temperatures. In a real crystal, the anisotropy will give preferred glide planes, and presumably the strip core would form in one of these planes.

[14] J. S. Koehler, in *Imperfections in Nearly Perfect Crystals*, edited by W. Shockley (John Wiley & Sons, Inc., New York, 1952), p. 197.

[15] A. Granato and K. Lucke, J. Appl. Phys. 27, 583 (1956).

[16] D. H. Niblett and J. Wilks, in *Advances in Physics*, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1960), Vol. 9, p. 1.

[17] A. S. Darling, J. Inst. Metals 85, 489 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

The disordered-core theory modifies to some extent the concepts of interaction between impurities and the dislocation in several ways, which we will now briefly enumerate. Impurities which condense onto the dislocation will not only form a Cottrell atmosphere, but will also become incorporated in the core, the ratio of the equilibrium concentration in the core to that immediately surrounding it being of the order of the segregation coefficient for normal solidification. A core containing impurities will have a higher thermodynamic potential, and hence in order to minimize this by dilution, a further increase of core width will occur. Motion of the dislocation away from the impurities will require the dissipation of a pinning energy, but this will be less than for the small-core dislocation normally considered. In addition, an oscillation of the dislocation with amplitude less than the dislocation width will not involve the same pinning energy, and the impurities will act as "soft" pinning points in that case. If a dislocation moves sufficiently slowly for impurities to follow, there will be a significant frictional force, the relaxation time of the process being of the order of the time for an impurity to diffuse across the core. All these points require further exposition, and it is intended to do this in a later paper.

A further effect of the disordered core will be seen in the "pipe diffusion" of defects. Diffusion coefficients for this process are high[18,19] and special structures, such as that of a hollow core, have been postulated to explain their magnitude. Queisser *et al.*[18] derive an activation energy for dislocation diffusion of phosphorous in silicon which is about 1.5 eV less than that of bulk diffusion. This difference was interpreted in terms of excess concentration of vacancies and impurities near the dislocation due to the formation of Cottrell atmospheres. However, the disordered-strip model shows that in the 1050 to 1200°C region where the measurements were made, there is appreciable broadening, which significantly lowers the Cottrell binding energy. Excess vacancy concentrations outside the core will therefore not be so large, and this mechanism of enhanced diffusion becomes less important. The alternative explanation is that the lower activation energy is associ- with the disordered-core material itself, which presumably would have an energy closer to that of the liquid state. Such a theory is being developed, taking into account possible changes of core size with impurity concentration in the core and the ratio of this to the concentration in the surrounding solid.

SUMMARY

The model presented here attempts to deal with the very complex problem of dislocation configuration and damping force by applying macroscopic thermodynamics to a microscopic situation. Although not strictly correct, the model has the advantage of involving a relatively simple conceptual picture of the dislocation at high temperatures, that of the thermally disordered core. The disordered material is considered to be quasi-liquid, so that appeal may be made to the concepts of latent heat of disorder, and of viscosity, in the calculations. "Melting" of the core relieves the stresses in the vicinity. It seems to be a reasonable assumption that a flat strip configuration does this most efficiently, in both edge and screw dislocations, although this has not been proven. A cylindrical disordered-core coaxial with the dislocation does not relieve stresses. The core width increases considerably as the melting point is approached, and will be greater than several atomic spacings over a significant temperature range in many crystals. Near the melting temperature, the surface energy of the order-disorder interface is a controlling factor in the size, and uncertainty in the value of the surface energy is reflected in an uncertainty in the width here.

Two separate contributions to the dislocation damping coefficient are derived: thermodynamic and viscous. The thermodynamic damping originates in the heat required to produce disorder in the forward edge of the moving dislocation and the entropy changes in transferring this heat from the solidifying trailing edge. It varies inversely as the temperature, but disappears when the dislocation narrows to only one or two atoms wide. Viscous damping decreases with temperature due to both the increase of dislocation width and decrease in viscosity of the disordered core. Both mechanisms appear to give damping coefficients of a significantly large order of magnitude; although good estimates of the viscosity of the disordered core are difficult to make.

Application of the theory to silicon shows that the damping will be large and will predominate near the melting point, while at lower temperatures, it is reasonable to suppose that viscous damping becomes important. The magnitudes are consistent with those derived from experiment. In the case of copper, the thermodynamic damping is expected to dominate the viscous over a wide temperature range, but the form of the temperature dependence deviates considerably from the experimentally observed variation. When the temperature becomes so low that the core is only two or three atoms wide, the disorder is better described as an assembly of kinks, and the more complex theories of kink generation and diffusion may then be used in place of the viscosity approximation.

A disordered core is expected to modify dislocation interactions with impurities, and to lead to a higher equilibrium concentration of impurities in the core, "soft" pinning effects, additional dislocation drag mechanisms, and enhanced impurity diffusion along the dislocation.

[18] H. J. Queisser, K. Hubner, and W. Shockley, Phys. Rev. 123, 1245 (1961).

[19] R. Tucker, A. Lasker, and R. Thomson, J. Appl. Phys. 34, 445 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

APPENDIX: DEPENDENCE OF STRESS-FIELD ENERGY ON CORE SIZE

1. Cylindrical Core

The dislocation core is considered to be a cylinder of radius ρ. We wish to calculate the stress-field energy per unit length, U_0, in the region outside a larger cylinder of radius ρ_0, and to obtain

$$\lim_{\rho_0 \to \rho} (\partial U_0/\partial \rho).$$

Consider a straight dislocation lying along the z axis and let σ^{ij} and e_{ij} be components of the stress and strain produced by this dislocation. Then

$$U_0 = \frac{1}{2}\int \sigma^{ij} e_{ij} dS, \tag{A1}$$

where the upper and lower indices denote contravariant and covariant components,[20] and the usual summation convention has been used. The integration is taken over the region of the xy plane between circles of radius ρ_0 and R (large). For an isotropic medium the stress and strain components are related by [21]

$$e_{ij} = \frac{1}{2G}\left[\sigma_{ij} - \frac{\nu}{\nu+1} g_{ij} g^{mn} \sigma_{mn}\right], \tag{A2}$$

where the g_{ij} are the components of the metric tensor.

The physical stress components[20] for a screw dislocation are given in Eq. (2.16) of Cottrell.[22] They are independent of ρ; hence U_0 will also be independent, and

$$\lim_{\rho_0 \to \rho} (\partial U_0/\partial \rho) = 0.$$

For an edge dislocation one cannot take the stress components directly from Eq. (2.10) of Cottrell. It is necessary to correct the components so that the stresses across the surface of the core vanish. Following the method outlined by Cottrell one obtains the corrected covariant stress components:

$$\begin{aligned}
\sigma_{rr} &= -\frac{Gb\sin\theta}{2\pi(1-\nu)r}\left[1 - \frac{\rho^2}{r^2}\right], \\
\sigma_{\theta\theta} &= -\frac{Gbr\sin\theta}{2\pi(1-\nu)}\left[1 + \frac{\rho^2}{r^2}\right], \\
\sigma_{\theta r} &= \sigma_{r\theta} = \frac{Gb\cos\theta}{2\pi(1-\nu)}\left[1 - \frac{\rho^2}{r^2}\right], \\
\sigma_{zz} &= \nu(\sigma_{rr} + (r^2)^{-1}\sigma_{\theta\theta}).
\end{aligned} \tag{A3}$$

Now, using the relation $\sigma^{ij} = g^{ik} g^{jl} \sigma_{kl}$ to obtain the contravariant stress components, and neglecting terms in $1/R^2$ and $1/R^4$, we derive

$$U_0 = \frac{Gb^2}{4(1-\nu)}\left[\ln\frac{R}{\rho_0} - \frac{\rho^2}{2(1-\nu)\rho_0^2} + \frac{\rho^4}{4(1-\nu)\rho_0^4}\right], \tag{A4}$$

and again

$$\lim_{\rho_0 \to \rho} (\partial U_0/\partial \rho) = 0.$$

We might note that $\lim_{\rho_0 \to \rho} U_0$ does not agree with the expression given by Cottrell [the right side of Eq. (2.22b) of Ref. 22]. Cottrell calculates the energy by considering a half-plane cut made in the unstressed material and determining the work done in displacing the two sides of the cut by the Burgers vector. However, this procedure ignores the work done in changing the shape of the core during the displacement, and it is apparent that this work is of the same order as the correction terms.

2. Flat Strip Core

Let the dislocation core be a strip of variable width 2ρ in the glide plane and fixed thickness q. Take a fixed boundary outside the core, width $2\rho_0$, and thickness q; U_0 is the elastic energy outside this boundary. We are interested in the case $q \ll \rho$ and hence, for simplicity, will calculate U_0 in the limit $q \to 0$. Suppose U' is the energy outside the strip of width 2ρ. As $q \to 0$, the volume between the two strips becomes vanishingly small and hence $U' \to U_0$. Since the limits involved in calculating the partial derivatives are independent of q, they will also become equal as $q \to 0$.

A convenient way of deriving U_0 is the use of the analogy between the magnetostatic field energy of a current and the elastic field energy of a dislocation.[23] The analogy is effected in this case by substituting the elastic modulus $G\beta$ for the permeability μ_0, and the Burgers vector b for the current I. We regard the strip as having a uniform distribution of Burgers vector between $-\rho$ and $+\rho$, and hence take a uniform current distribution for the analogous case. The magnetostatic energy is related to the inductance L by

$$U_m = \tfrac{1}{2} L I^2,$$

while L is readily obtained from the formulas of Oberhettinger and Magnus[24] for the characteristic impedance of a transmission line consisting of a flat strip surrounded by a large outer conductor. If the outer conductor is sufficiently large, a good approximation is

$$L = (\mu_0/2\pi)\ln(kR/\rho),$$

[20] See, for example, G. E. Hay, *Vector and Tensor Analysis* (Dover Publications Inc., New York, 1953), Chap VI.

[21] This can be obtained by generalizing Eq. (12.3) of Ref. 24, which is given in Cartesian coordinates. The m in this equation is the reciprocal of Poisson's ratio ($m = 1/\nu$).

[22] A. H. Cottrell, *Dislocations and Plastic Flow in Crystals* (Oxford University Press, London, 1953).

[23] R. de Wit, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1960), Vol. 10, p. 249.

[24] F. Oberhettinger und W. Magnus, *Anwendung der Elliptischen Funclionen in Physik und Technik* (Springer-Verlag, Berlin, 1949), p. 63.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

where k is of the order of 2.5 and depends on the geometry of the outer conductor, and mks units are used. Therefore

$$U_0=(G\beta b^2/4\pi)\ln(kR/\rho)$$

and

$$\partial U_0/\partial\rho=-(G\beta b^2/4\pi\rho).$$

The same expression is obtained by an approximation in which the material within a circle of radius ρ is taken to be uniformly stressed, in tension above and in compression below the strip, while outside the circle it is stressed just as if the strip were of infinitesimal width. This relatively simple stress distribution appears to be a good approximation to the exact situation.

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Surface Effect in Secondary and Photoelectric Emission

I. ADAWI
Battelle Memorial Institute, Columbus, Ohio

(Received 10 January 1964; revised manuscript received 28 February 1964)

The secondary emission yield of metals by the surface effect δ is expressed in terms of the surface photoelectric yield $y(\omega)$ for a radiation of frequency ω and angle of incidence of $\cos^{-1}\frac{1}{2}(5^{\frac{1}{2}}-1)$ which is about 52°. It is shown that $\delta\sim(2\pi\alpha E_p)^{-1}\int_{\omega_1}^{\omega_2} y(\omega)d\omega/\omega$, where E_p is the primary energy in atomic units, $\alpha=1/137$, ω_1 is the threshold frequency, and ω_2 depends on the energy of the primary and may be replaced by ∞. For a square-well potential model for a metal, $\delta\sim10^{-3}/E_p$ with a relative error of order $(E_F/E_p)\ln(E_p/E_F)$, where E_F is the Fermi energy.

1. INTRODUCTION

THE purpose of this paper is to establish a general and simple relation between the surface effect in secondary electron emission[1] (SSE) from metals and the surface photoelectric effect[2,3] (SPE), and to use this relation to show clearly why the SSE is so small that it can be neglected in explaining the experimental facts. Such a relation is of interest for its own sake, and in addition a new examination of the problem is desirable since most of the published papers on the SSE, are incorrect.[4,5]

We shall use for convenience Hartree's atomic units in which $\hbar$, the electron mass m, the Bohr radius a_B, and the electron charge e are unity, and the speed of light c is 137. As was done in discussing the surface photoelectric effect in I, we treat the conduction electrons as independent noninteracting particles. The motion of a single electron in the y and z directions is free and can be described by the plane-wave $L^{-1}\exp i(k_y y+k_z z)$ obeying cyclic boundary conditions and normalized to unity in a square area of side L. The x motion is bound by a general surface potential $V(x)$ which is the same for all electrons and varies only in the direction x which is normal to the metal surface. This motion is described by the wave function $L^{-1/2}\phi_0(x)$ normalized to unity. The length L is the thickness of the metal plate which extends from $x=-L$ to $x=0$. The function $\phi_0(x)$ satisfies the wave equation,

$$H_1\phi_0=E_0\phi_0, \tag{1.1}$$

where

$$H_1(x)=-\tfrac{1}{2}\partial^2/\partial x^2+V(x). \tag{1.2}$$

Inside the metal, $V(x)$ is a constant equal to $-V_0$, and we can write $\phi_0(x)=2^{1/2}\sin(k_x x+\gamma)$ where γ is a phase factor depending on $V(x)$, and $\frac{1}{2}k_x^2=E_0+V_0$. The potential $V(x)$ rises to zero in the surface barrier regions near $x=0$ and $x=-L$ in a distance much less than L, and $\phi_0(x)$ behaves as $\exp(-px)$ for large x, where $-\frac{1}{2}p^2=E_0$. We ignore thermal effects and assume that all energy states below the Fermi energy E_F are occupied, and all energy states above the Fermi energy are empty. The conduction electrons in this model assume k values which fill "a Fermi hemisphere" given by $k^2=k_F^2$ and $k_z>0$, where k_F is the Fermi momentum,

[1] For a review of secondary emission see O. Hachenberg and W. Brauer, in *Advances in Electronics and Electron Physics* (Academic Press Inc., New York, 1959), Vol. XI, p. 413; A. J. Dekker, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1958), Vol. 6, p. 251.

[2] K. Mitchell, Proc. Roy. Soc. (London) **A146**, 442 (1934).

[3] I. Adawi, Phys. Rev. **134**, A788 (1964). This paper will be denoted by I.

[4] A. Viatskin, Zh. Eksperim. i Teor. Fiz. **9**, 826 (1939) treated a semi-infinite square-well potential model. The basic formulation is correct but the final integrations and conclusions are obscure and incorrect.

[5] W. Brauer and W. Klose, Ann. Physik **19**, 116 (1956). This paper has been assumed correct in the two review articles cited in Ref. 1, but it contains unfortunately basic errors. They treat a finite square well of width $2a$ in the limit that $a\to\infty$. The correct final state which should be used in the transition matrix element is the function v^* used here. Using the notation and Eq. (4.5) of I we have that the incoming wave v is given by $v=\frac{1}{2}(\phi^s/\Lambda_s^*+\phi^a/\Lambda_a^*)$. When the correct limiting procedure is applied as $a\to\infty$, the results of the finite and the semi-infinite square well become identical as has been discussed in general in I. With this in mind, and for a primary electron incident normal to the metal surface, none of the four delta functions obtained by Brauer and Klose and on which essentially all their discussion is based should arise; and the effect is precisely determined by terms similar to those they ignored. The yield by the surface effect in secondary emission as in photoelectric emission is independent of the dimensions of the model analyzed, and there is no need to introduce an *ad hoc* depth d_s for calculating the effect.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

i.e., $E_F=\frac{1}{2}k_F^2$. The number of electrons per unit volume in the interval d^3k is given by $4d^3k/(8\pi^3)$, where the factor 4 is contributed by the usual factor 2 for spin and an additional factor 2 to take account of the distribution of the bound states of the x motion in the continuum limit[3] (large $V_0^{1/2}L$).

We consider a primary electron of momentum $\mathbf{K}$ and energy $E_p=\frac{1}{2}K^2$ incident *normal* to the metal surface, namely, $\mathbf{K}=(-K,0,0)$. The mutual Coulomb repulsion between the primary electron and the conduction electrons bound by the surface potential offers a *direct* mechanism for a secondary electron emission from the bombarded metal surface. It is with this process that we are here concerned. We do not concern ourselves with those electrons which are left excited after colliding with the primary, and which diffuse in the metal, and appear later as secondaries. We hardly need to emphasize that the effect we are discussing exists only by virtue of the binding of the conduction electrons to the surface potential, since it is well known that no secondaries can be *directly* emitted from a metal surface bombarded with *normally* incident primaries if the electron gas is completely free.[6]

In Sec. 2 we discuss the scattering problem of the system consisting of a primary electron and a conduction electron, hereafter referred to as secondary. We shall consider only fast primaries (nonrelativistic) whose energy is much greater than the Fermi energy. Indeed, the situations of most practical interest[1] are those where the primary energy E_p is of order 10 (272 eV) and E_p/E_F of order 100. We are then justified in treating the Coulomb field of the two electrons as a perturbation and neglecting the effect of the surface potential on the primary wave functions. We shall, however, treat exactly the effect of the surface potential on the motion of the secondary in the final state which leads to the use of the well-known "incoming wave"[3,7] in the transition matrix element. The two electrons will be treated as distinguishable particles described by a product wave function, since we shall see that the main contributions to the SSE come from small-angle scattering for which exchange effects are not important.

In Sec. 3 we use the small-angle scattering approximation and derive the main result of the paper which expresses the secondary emission yield by the surface effect as an integral over the surface photoelectric yield. In Sec. 4 we illustrate this result by treating two examples. In the first example we use experimental measurements on the SPE. In the second example we apply our formulas to the previously treated[4,5] square-well potential model for the metal, and obtain the yield by the SSE which turns out to be in agreement with Baroody's unpublished result.[8]

[6] H. Fröhlich, Ann. Physik 5, 13, 229 (1932).

[7] See G. Breit and H. A. Bethe, Phys. Rev. 93, 888 (1954); I. Adawi, Am. J. Phys. 32, 211 (1964).

[8] E. M. Baroody (unpublished report); and abstract in Phys. Rev. 92, 843 (1953).

2. FORMULATION

The unperturbed Hamiltonian of the system H_0 consists of three commuting parts: H_1 defined in (1.2), H_2 the kinetic energy of the secondary for the y and z motions, and H_3 the kinetic energy of the primary. Thus,

$$H_0=H_1+H_2+H_3, \tag{2.1}$$

$$H_2=-\tfrac{1}{2}(\partial^2/\partial y^2+\partial^2/\partial z^2), \tag{2.2}$$

$$H_3=-\tfrac{1}{2}\nabla_{\mathbf{R}}^2, \tag{2.3}$$

where $\mathbf{r}=(x,y,z)$ is the position of the secondary and $\mathbf{R}$ is the position of the primary. The perturbation potential H' is given by the Coulomb field,

$$H'=1/|\mathbf{R}-\mathbf{r}|. \tag{2.4}$$

We consider as in Sec. 1 a metal plate defined by $-L<x<0$, and discuss the secondary emission from a square region of area L^2 of the surface $x\approx 0$ which is bombarded by a uniform primary beam moving in the $-x$ direction. The interaction volume is L^3 and we normalize the primary wave function to unity in this volume. Thus the initial primary state is $L^{-3/2}\exp(i\mathbf{K}\cdot\mathbf{R})$ with $\mathbf{K}=(-K,0,0)$. The initial state of the secondary has been defined, and we take for the initial state of the system the wave function ψ_0 defined by

$$\begin{aligned} &H_0\psi_0=\mathcal{E}\psi_0,\\ &\psi_0=L^{-3}\exp i(\mathbf{K}\cdot\mathbf{R}+\mathbf{k}_\perp\cdot\mathbf{r})\phi_0(x),\\ &\mathcal{E}=E_0+\tfrac{1}{2}(k_\perp^2+K^2), \end{aligned} \tag{2.5}$$

where for any vector $\mathbf{s}$, $\mathbf{s}_\perp\equiv(0,s_y,s_z)$.

The solution ψ^+ of the perturbed problem for outgoing waves satisfies the scattering integral equation,[9,10]

$$\psi^+=\psi_0+[1/(\mathcal{E}-H_0+i\epsilon)]H'\psi^+. \tag{2.6}$$

If we write

$$\psi^+=\psi_0+\psi_s, \tag{2.7}$$

where ψ_s is the scattered wave, we obtain to first order in H' that

$$\psi_s=[1/(\mathcal{E}-H_0+i\epsilon)]H'\psi_0. \tag{2.8}$$

We expand H' in terms of the eigenfunctions of the primary and write the usual Fourier series,

$$H'=1/|\mathbf{R}-\mathbf{r}|=\textstyle\sum_\kappa(4\pi/\kappa^2)L^{-3}\exp[i\boldsymbol{\kappa}\cdot(\mathbf{r}-\mathbf{R})], \tag{2.9}$$

in which the components of $\boldsymbol{\kappa}$ are given by $(2\pi/L)$ times an integer which is also the case for k_y and k_z.

If in (2.8) we substitute for H' and ψ_0 the expressions given by (2.9) and (2.5) and observe that the relation

$$\begin{aligned} (\mathcal{E}-H_0+i\epsilon)^{-1}&\exp i(\mathbf{K}'\cdot\mathbf{R}+\mathbf{k}_\perp'\cdot\mathbf{r})\\ &=[\exp i(\mathbf{K}'\cdot\mathbf{R}+\mathbf{k}_\perp'\cdot\mathbf{r})]\\ &\quad\times[\mathcal{E}-H_1-\tfrac{1}{2}(K'^2+k_\perp^2)]^{-1}, \end{aligned} \tag{2.10}$$

[9] B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 469 (1950).

[10] M. Gell-Mann and M. L. Goldberger, Phys. Rev. 91, 398 (1953).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

holds for a general $\mathbf{K}'$ and $\mathbf{k}'$, we obtain

$$\psi_s=\sum_\kappa \exp\{i[(\mathbf{K}-\boldsymbol{\kappa})\cdot\mathbf{R}+(\mathbf{k}_\perp+\boldsymbol{\kappa}_\perp)\cdot\mathbf{r}]\} \times(4\pi/\kappa^2)L^{-6}\phi(\boldsymbol{\kappa},x), \quad (2.11)$$

$$\phi(\boldsymbol{\kappa},x)=\int G^+(E';x,x')e^{i\kappa_x x'}\phi_0(x')dx', \quad (2.12)$$

$$G^+(E')\equiv(E'-H_1+i\epsilon)^{-1}, \quad (2.13)$$

$$E'\equiv\tfrac{1}{2}q^2\equiv E_0+\omega, \quad (2.14)$$

$$\omega=-\kappa_x K-\tfrac{1}{2}\kappa_x^2-\kappa_\perp^2-\mathbf{k}_\perp\cdot\boldsymbol{\kappa}_\perp, \quad (2.15)$$

which reduces the problem to a one-dimensional problem, since the Green's function $G^+(E')$ contains only the Hamiltonian of the x motion.

To obtain the secondary current we examine the asymptotic behavior of ψ_s for positive large x. For large x and finite x' we have as in I that,

$$G^+(E';x,x')\sim(iq)^{-1}e^{iqx}v(x'), \quad (2.16)$$

where $v(x)$ is the "incoming wave" defined as the *incident wave* $\exp(-iqx)$ *plus the reflected and transmitted waves due to the potential* $V(x)$. From (2.16) and (2.12)

$$\phi(\boldsymbol{\kappa},x)\sim(iq)^{-1}Me^{iqx}, \quad (2.17)$$

$$M\equiv\int v(x)e^{i\kappa_x x}\phi_0(x)dx\equiv\langle v^*|\exp(i\kappa_x x)|\phi_0\rangle. \quad (2.18)$$

From the cylindrical symmetry of the problem, the total secondary particle current is in the x direction, and we need to calculate, therefore, the current element dI_x contributed by ψ_s and threading the area L^2. Using (2.17) in (2.11) and the usual formula,

$$dI_x(x)=\frac{1}{2i}\int\left[\psi_s^*\frac{\partial\psi_s}{\partial x}-\psi_s\frac{\partial\psi_s^*}{\partial x}\right]d^3Rdydz, \quad (2.19)$$

we obtain

$$dI_x=\sum_\kappa 16\pi^2\kappa^{-4}|M|^2q^{-1}L^{-7}. \quad (2.20)$$

Notice that in (2.19) the y and z integrations give L^2, the $\mathbf{R}$ integration gives L^3, and the orthogonality of the primary final states eliminates any interference terms between different values of $\boldsymbol{\kappa}$. The total current I_x is obtained by summing (2.20) over the conduction electrons in the volume L^3, and if we replace the summations by integrations we obtain

$$I_x=L^{-1}\int(d^3k/2\pi^3)\int(2/\pi)d^3\kappa(\kappa^{-4}|M|^2q^{-1}), \quad (2.21)$$

where the $\mathbf{k}$ integration is extended over the Fermi hemisphere and the $\boldsymbol{\kappa}$ integration is restricted by the condition that q is real.

If the normally incident primary is to lose energy to the secondary, κ_x *must be negative*. Furthermore, ω of (2.15) must exceed the metal work function ω_1 before any secondaries can be emitted *directly* by the present mechanism. This implies that some screening of the Coulomb field is automatically included in the model and we see from (2.15) that, to a good approximation, all $\kappa<\omega_1/K$ do not contribute.[11]

The SSE yield δ, which is the number of secondaries emitted per incident primary, is obtained by dividing I_x by the primary current K/L, namely,

$$\delta=(2/\pi K)\int(d^3k/2\pi^3)\int d^3\kappa(\kappa^{-4}q^{-1}|M|^2). \quad (2.22)$$

To show clearly that δ is independent of L as $L\to\infty$, we must investigate the nature of M in this limit. It was shown in I that in discussing electron emission from the surface $x=0$, a finite plate model with two surface barriers, one at $x=0$ and the other at $x=-L$, reduces in the limit $L\to\infty$ to a semi-infinite model with one barrier at $x=0$. The limiting "incoming wave" $v(x)$, which must be used in (2.18), is a solution of the wave equation for this limiting potential for $x>-L$, and is zero for $x<-L$. Thus inside the metal $v(x)\propto\exp(-ik_x'x)$ where a cutoff is implied to render $v(x)$ zero as $x\to\infty$, and where $k_x'^2=q^2+2V_0=k_x^2+2\omega$. The integrand in (2.18) for $x<0$ is of the form

$$\exp i(-k_x'\pm k_x+\kappa_x)x$$

for which the phase cannot vanish, since $k_x'>k_x$ and $\kappa_x<0$, and hence the integral cannot display a delta-function-type singularity which, if it were present, would bring a length[12] L in $|M|^2$. Obviously, the integration in (2.18) over x between $x=0$ and $x=\infty$ is convergent since $\phi_0(x)$ is a damped wave for large x. We conclude that except for thin films, to which still the present formulation can easily be adapted, $|M|^2$, and δ are independent of L, and we can set formally $L=1$ in the preceding equations.

For completeness, we give now the results of the time proportional transitions method in discussing the scattering problem. A detailed derivation can be found in Gell-Mann and Goldberger,[10] or obtained by other means. The transition from the initial state ψ_0 of (2.5) to the plane-wave final state ψ_f of energy $\mathcal{E}'$, where $\psi_f=\exp i(\mathbf{K}'\cdot\mathbf{R}+\mathbf{k}'\cdot\mathbf{r})$, $\mathbf{K}'=\mathbf{K}-\boldsymbol{\kappa}$ and $\mathbf{k}'=(q,k_y',k_z')$, is determined by the transition matrix element T_{f0}

---

[11] For a uniform electron gas screening sets in for momenta $<\omega_p/K$ where ω_p is the plasma frequency, see, e.g., H. A. Kramers, Physica 13, 401 (1947), J. Lindhard, Kgl. Danske Viedenskab. Selskab, Mat. Fys. Medd. 28, 8 (1954). In the surface problem under consideration, the electron density, plasma frequency, and dielectric constant are variable in the region of interest, and it is not deemed worthwhile to discuss screening beyond what naturally occurs in the problem.

[12] Notice that as $L\to\infty$, $2\pi\delta(k)=\int_{-1/2L}^{1/2L}dx\exp(ikx)$, and $2\pi|\delta(k)|^2$ is interpreted as $L\delta(k)$. If M had a delta function behavior the sum over k_x would have to be restricted to a length d_s and not L, where d_s is a characteristic escape depth for the secondary, but then we would be really discussing a volume effect!

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

which is given by

$$T_{f0}=\langle\psi_f|H'+V|\psi^+\rangle. \tag{2.23}$$

After some algebra (2.23) gives to first order in H' the result,

$$T_{f0}=\langle\psi_f^-|H'|\psi_0\rangle, \tag{2.24}$$

where $\psi_f^-=v^*(x)\exp i(\mathbf{K}'\cdot\mathbf{R}+\mathbf{k}_\perp'\cdot\mathbf{r})$, the only difference between (2.24) and the simple version of the Born approximation is the replacement of $\exp(iqx)$ by $v^*(x)$. The $\mathbf{R}$ integration in (2.24) gives $4\pi\kappa^{-2}$, and the y and z integrations give Kronecker deltas, since momentum is conserved for these directions, and we have

$$T_{f0}=4\pi\kappa^{-2}M\delta_{k_y',k_y+\kappa_y}\delta_{k_z',k_z+\kappa_z}. \tag{2.25}$$

By summing the transition rate, $2\pi|T_{f0}|^2\delta(\mathcal{E}'-\mathcal{E})$, over k_y' and k_z' holding $\boldsymbol{\kappa}$ fixed, and integrating over q we obtain $16\pi^2\kappa^{-4}|M|^2q^{-1}$, which when summed over $\boldsymbol{\kappa}$ and $\mathbf{k}$ gives exactly Eq. (2.21) for the total current (with $L=1$).

3. RESULT

The yield δ of (2.22) is controlled mainly by the Coulomb scattering which is proportional to κ^{-4}, and by the matrix element M which is the Fourier transform of $v\phi_0$, where v and ϕ_0 are eigenfunctions of H_1 whose energy difference is ω. The problem is basically the same as the ionization of atoms by fast electrons for which the physical principles are well established.[13] We shall here use the energy k_F^2 as a characteristic energy to obtain order of magnitude estimates, in much the same way as the binding energy is used in atomic collisions. For $\kappa_x=0$, $M=0$, and for small κ_x we write $\exp(i\kappa_x x)\sim1+i\kappa_x x$, and we have the well-known dipole approximation,

$$M\approx i\kappa_x\langle v^*|x|\phi_0\rangle. \tag{3.1}$$

From the commutation relations, $[x,H_1]=\partial/\partial x\equiv D$, and $[D,H_1]=\partial V/\partial x$, we can rewrite[3] (3.1) as

$$\begin{aligned} M&\approx -i\kappa_x\omega^{-1}M_1,\\ M_1&\equiv\langle v^*|D|\phi_0\rangle\equiv-\omega^{-1}\langle v^*|V'|\phi_0\rangle. \end{aligned} \tag{3.2}$$

For $\kappa_x\gg k_F$ the binding effect of the potential $V(x)$ should become negligible, and the conduction electrons can be treated as free electrons. Since we know that in such a situation no secondaries can be emitted,[6] we conclude that M must decrease very rapidly as κ_x becomes large. This coupled with the fact that the Coulomb field strongly favors small momenta transfer, leads to the conclusion that the major contributions to the yield come from those values of κ for which $\kappa_x\ll k_F$ and $\omega\sim k_F^2$. The dipole approximation (3.1) applies and there is no need to consider large values of κ_x.

The expression (2.15) for ω will now be simplied. To estimate various terms, we recall that for a fractional energy loss Δ by the primary and for a scattering angle Θ we have $K'\approx(1-\frac{1}{2}\Delta)K$, $-\kappa_x=\frac{1}{2}(\Delta+\Theta^2)K$, and $\kappa_\perp\approx\Theta K$. The term $\frac{1}{2}\kappa_x^2$ can certainly be neglected. The term $\mathbf{k}_\perp\cdot\boldsymbol{\kappa}_\perp$ is zero on the average, and it is small compared to $-\kappa_x K$ for $\Theta^2\ll\Delta\sim k_F^2/K^2$ or $\Theta<k_F/K$. For $\Theta\gg k_F/K$, $\kappa_\perp^2\gg\boldsymbol{\kappa}_\perp\cdot\mathbf{k}_\perp$. We can, therefore, neglect the term $\boldsymbol{\kappa}_\perp\cdot\mathbf{k}_\perp$ without committing a serious error. Equation (2.15) now takes the approximate form

$$\omega\approx-\kappa_x K-\kappa_\perp^2, \tag{3.3}$$

whose accuracy will be later discussed.

By using (3.2) and (3.3) in (2.22) we can hold $\boldsymbol{\kappa}$ fixed and integrate over $\mathbf{k}$, and we can write

$$\delta=(\pi^2\alpha K)^{-1}\int d^3\kappa\,\kappa_x^2\omega^{-1}y(\omega), \tag{3.4}$$

where, by definition,

$$y(\omega)=2\pi\alpha\omega^{-1}\int(d^3k/2\pi^3)q^{-1}|M_1|^2, \tag{3.5}$$

and $\alpha=1/137$. To identify $y(\omega)$ we shall use the results of I on the first-order SPE. For a monochromatic radiation beam of angular frequency ω and polarization ($\mathbf{E}$ vector) in the plane of incidence, we find by summing Eq. (2.15a) of I over $\mathbf{k}$ that the photoelectric current density I_1 is given by

$$I_1=137ny(\omega)\sin^2\theta, \tag{3.6}$$

where n is the number of photons/unit volume in the incident beam, and θ is the angle of incidence. By dividing I_1 by $nc\cos\theta$ which is the normal component of the photon current density, we obtain the SPE yield $Y(\omega,\theta)$, namely,

$$Y(\omega,\theta)=y(\omega)(\sin^2\theta/\cos\theta). \tag{3.7}$$

Thus, $y(\omega)$ is the SPE yield for $\theta=\cos^{-1}\frac{1}{2}[(5)^{1/2}-1]=51°50'$.

The azimuthal integration in (3.4) gives 2π. The remaining double integral is written using the new variables, $\omega=-\kappa_x K-\kappa_\perp^2$ as in (3.3) and $u=\kappa_\perp^2$, and we have

$$\delta=(2\pi\alpha E_p)^{-1}\int_{\omega_1}^{\omega_2}(d\omega/\omega)y(\omega)F(\omega), \tag{3.8}$$

where

$$F(\omega)=2E_p\int_0^{u_0}du\left(1+E_p\frac{\partial}{\partial E_p}\right)[(u+\omega)^2+2E_pu]^{-1}. \tag{3.9}$$

Here, $\omega_2=\kappa_{x0}K$, where κ_{x0} is a maximum value of $-\kappa_x$ consistent with the approximation $-\kappa_x\ll k_F$, and u_0

[13] H. A. Bethe, Ann. Physik 5, 325 (1930). See also N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Oxford University Press, New York, 1952), 2nd ed., Chap. XI; L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, translated by J.B. Sykes and J. S. Bell (Addison Wesley Publishing Company, Inc., Reading, Massachusetts, 1958), Chap. XV.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

$=\omega_2-\omega$. Performing the integration in (3.9), we obtain

$$F(\omega)=\frac{\omega E_p^{1/2}}{(E_p+2\omega)^{3/2}}\ln\frac{u-u_1}{u-u_2}-\frac{2E_p[(E_p+\omega)u+\omega^2]}{(E_p+2\omega)[u^2+2(E_p+\omega)u+\omega^2]}\Bigg|_0^{u_0},$$

where

$$u_{1,2}\equiv-(E_p+\omega)\pm(E_p^2+2E_p\omega)^{1/2}. \quad (3.10)$$

For $\omega^2/E_p\ll u_0\ll E_p$, Eq. (3.10) gives

$$F(\omega)\approx 1+(\omega/E_p)\ln(1+2u_0E_p\omega^{-2}). \quad (3.11)$$

It is evident that $F(\omega)$ is a slowly varying function of ω and u_0 and hence the SPE yield in (3.8) controls the SSE yield. It is well known that[2,3] $y(\omega)$ reaches a peak for $\omega\sim k_F^2$ and drops rapidly with increasing frequency. If we take $\kappa_{z0}\sim k_F/4$ we see that $\omega_2\gg k_F^2$ for $K/k_F\gg 4$ and ω_2 may be replaced by ∞ in (3.8). This shows that the important values of ω are of order k_F^2 and that the emitted secondaries are of low energy. If $u_0\ll\omega^2/E_p$, (3.10) shows that $F(\omega)$ is small, but then $y(\omega)$ is small since $\omega_2\gg k_F^2$ and these values of ω are not decisive in determining δ. We set $u_0=\omega_2$ in (3.11) and neglect the one in the logarithmic term, and substitute the result in (3.8) to obtain

$$\delta\approx(2\pi\alpha E_p)^{-1}\int_{\omega_1}^{\infty}(d\omega/\omega)y(\omega)\times[1+(\omega/E_p)\ln(2\omega_2E_p\omega^{-2})]. \quad (3.12)$$

To simplify (3.12) we treat the logarithmic term as a constant in which we set $\omega_2=\beta_1k_FK$ and $\omega=\omega_{\max}$, where β_1 is a small fraction and $\omega_{\max}$ is the angular frequency for which the SPE yield $y(\omega)$ is maximum. We can write $\omega_{\max}=\beta_2k_F^2$ where β_2 is of order unity, and denote β_1/β_2^2 by a new parameter β. We obtain the result,

$$\delta\approx(2\pi\alpha E_p)^{-1}\int_{\omega_1}^{\infty}(d\omega/\omega)y(\omega)\times[1+(\omega/E_p)\ln\beta(E_p/E_F)^{3/2}]. \quad (3.13)$$

The leading term in (3.13) would have been obtained if we neglected the term κ_1^2 in (3.3). The logarithmic term gives a correction term of order $\frac{3}{2}(E_F/E_p)\times\ln(E_p/E_F)$ which is about 25% for $E_p/E_F\sim 20$.

Let us now return to the approximations made:

(*a*) $\cdot\mathbf{k}_1\cdot\boldsymbol{\kappa}_1$ *term:* This is the most bothersome term in the calculation and was neglected in (3.3). We can estimate the error by including this term in (3.9) and averaging the denominator over $\mathbf{k}$ to obtain $(u+\omega)^2+2u(E_p+E_F/5)$ for the new denominator. The integration can be carried out and we obtain the correction term, $(E_F/5E_p)\ln\beta(E_p/E_F)^{3/2}$, to be added inside the bracket of (3.13). Since this correction is smaller than the second term retained in (3.13), we can ignore it.

(*b*) *Large* κ_x: If in (3.10) we set $u_0=E_p-\omega$ which would allow for the primary to lose all its energy we obtain (for $\omega\ll E_p$) that

$$F(\omega, E_p-\omega)\approx(4/3)+(\omega/E_p)\ln(4E_p^2/3\omega^2). \quad (3.14)$$

Comparing (3.14) and (3.11) we see how little the large momenta transfers contribute to the effect. Since the dipole approximation of (3.2) is expected to overestimate M for large κ_x, Eq. (3.14) cannot be trusted and will be discarded.

(*c*) $-\frac{1}{2}\kappa_x^2$ *term:* This term can easily be included in (3.3). The resulting integral is similar to (3.9) and can be evaluated exactly. We find that $F(\omega)$ of (3.11) remains the same (except that 1 is replaced by $\frac{1}{2}$ inside the logarithm) and (3.13) is not altered. Equation (3.14), however, now takes the form,

$$F(\omega, E_p-\omega)\approx(5/3)+(\omega/E_p)\ln(8E_p^2/3\omega^2), \quad (3.15)$$

to which the previous remarks apply.

Thus our approximations are good, and we conclude that the simple formula,

$$\delta\sim(2\pi\alpha E_p)^{-1}\int_{\omega_1}^{\infty}(d\omega/\omega)y(\omega), \quad (3.16)$$

is quite satisfactory, and that (3.13) contains the major corrections to the SSE yield δ.

4. EXAMPLES

A. Experimental

To obtain an estimate of (3.16), let us use the experimental yield curves $Y(\omega,\theta)$ for Na and K which are reproduced in Weissler's review article.[14] From the relation (3.7) we obtain $y(\omega)$. By a *crude* numerical integration we find that the yield integrals, $\int y(\omega)d\omega/\omega$ are $\sim 6.7\times10^{-4}$ and 4×10^{-5} for Na and K, respectively, which lead to the values $\delta\sim 1.5\times10^{-2}/E_p$ and $0.9\times10^{-3}/E_p$, respectively. This shows clearly that the effect is entirely negligible, since the observed values[1] of δ are of order unity for $E_p\sim 10$.

B. Theoretical

A theoretical calculation of δ is possible for a square-well potential model. Let $V(x)=-V_0$ for $x<0$, and $V(x)=0$, for $x>0$. The wave functions v and ϕ_0 can be

[14] G. L. Weissler, *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1956), Vol. 21, p. 351. For Na we use the measurements of Maurer at $\theta=60°$, and for K the measurements of Suhrmann and Theissing at $\theta=65°$. The yield curve for K has to be cut off, since a true SPE must asymptotically drop rapidly with frequency. For square well potential, for example, $y(\omega)\propto\omega^{-7/2}$ as $\omega\to\infty$ as can be deduced from Eqs. (4.4) and (4.5).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

written down easily as in I, namely,

$$v(x)=\left(\frac{2q}{q+k_x'}e^{-ik_x'x},\ e^{-iqx}+\frac{q-k_x'}{q+k_x'}e^{iqx}\right),$$
$$\phi_0(x)=2^{-1/2}\left(\frac{k_x-ip}{k_x+ip}e^{-ikxx}+e^{ikxx},\ \frac{2k_x}{k_x+ip}e^{-px}\right) \quad (4.1)$$
$$\text{for } x<0 \quad \text{and} \quad x>0,$$

where the propagation constants have been previously defined. By using (4.1) in (2.18) we obtain after some lengthy algebra:

$$|M|^2=32\kappa_x^2q^2k_x^2(k_x'-q)(k_x'+q)^{-1}/R,$$
$$R\equiv(2\omega-2\kappa_xk_x'+\kappa_x^2)^2[4\omega^2+\kappa_x^2(\kappa_x^2+4p^2-4\omega)]. \quad (4.2)$$

We see that if $\frac{1}{2}\kappa_x^2\ll\omega$, $R\approx16\omega^4$ and

$$|M|^2\approx2\kappa_x^2q^2k_x^2(k_x'-q)(k_x'+q)^{-1}\omega^{-4}, \quad (4.3)$$

which is precisely the dipole approximation of (3.2) with which we could have started. The advantage of giving the exact $|M|^2$ is to show how good are the approximations of Sec. 3. It is evident from (4.2) that for a finite ω and large κ_x, $|M|^2\propto\kappa_x^{-6}$ which justifies the neglect of large values of κ_x.

It follows from Eqs. (4.3), (3.2), and (3.5) above, or from Eqs. (3.9) and (3.10) of I, that

$$y(\omega)=(16\alpha/\pi)(\omega/E_F)^{-3}J_1(\omega), \quad (4.4)$$

$$J_1(\omega)=\frac{1}{2\eta}\int_{\epsilon+\Omega-\eta>0}^{\epsilon=1}(1-\epsilon)(\epsilon+\Omega-\eta)^{1/2}\times\epsilon^{1/2}[(\epsilon+\Omega)^{1/2}-(\epsilon+\Omega-\eta)^{1/2}]^2d\epsilon, \quad (4.5)$$

where $\Omega=\omega/E_F$, $\epsilon=(V_0+E_0)/E_F$, and $\eta=V_0/E_F$.

If we substitute (4.4) and (4.5) in (3.13), define the parameter λ by

$$\lambda=(E_F/E_p)\ln\beta(E_p/E_F)^{3/2}, \quad (4.6)$$

and introduce the integration variables ϵ and ζ where $\zeta=\epsilon+\Omega-\eta$, we have

$$\delta=\frac{4}{\pi^2E_p\eta}\int_0^\infty d\zeta[(\zeta+\eta)^{1/2}-\zeta^{1/2}]^2\times\zeta^{1/2}\int_0^1(1-\epsilon)\epsilon^{1/2}S(\lambda,\zeta,\epsilon)d\epsilon, \quad (4.7)$$

where

$$S(\lambda,\zeta,\epsilon)=(\eta+\zeta-\epsilon)^{-4}+\lambda(\eta+\zeta-\epsilon)^{-3}. \quad (4.8)$$

Expanding S in a power series of ϵ and integrating term by term we obtain

$$\delta=\frac{16}{\pi^2E_p}\eta^{-5/2}\left\{\frac{1}{15}\eta C_{-1}\lambda+\sum_{n=0}^{\infty}\binom{-4}{n}(-\eta)^{-n}\frac{1}{(2n+5)}\times\left[\frac{1}{2n+3}+\frac{3\lambda}{(n+1)(2n+7)}\right]C_n\right\}, \quad (4.9)$$

where

$$C_n=\int_0^\infty\zeta^{1/2}[(\zeta+1)^{1/2}-\zeta^{1/2}]^2(\zeta+1)^{-(4+n)}d\zeta$$
$$=(-1)^{n+1}\binom{1/2}{n+2}\frac{2n+9}{2n+6}\pi-\frac{8}{(2n+5)(2n+3)}, \quad (4.10)$$

$C_{-1}=7\pi/8-8/3=0.0822$, $C_0=(3\pi/16)-8/15=0.0557$,

$C_1=(11\pi/128)-8/35=0.0414$.

The leading term in (4.9) is $\delta\sim0.06(\pi^2E_p\eta^{5/2})^{-1}$ which is of order $10^{-3}/E_p$ since η is about 2. The relative error committed by neglecting terms in λ is of order $[\eta(C_{-1}/C_0)+9/7]\lambda\sim4\lambda$, which could amount to 25% for $E_p/E_F\sim100$ (if we take $\beta=1$).

This agrees with Baroody[8] who obtained the first two terms of (4.9) with $\lambda=0$, by using the same method as did Viatskin, but making an independent evaluation of the final integral. It also agrees surprisingly well with the estimate obtained above from the measurements on the surface photoelectric effect of potassium.

Note added in proof. Equation (3.16) predicts that the energy loss of the primary (essentially ω) by SSE is nearly peaked at the energy value for which $y(\omega)/\omega$ is maximum, and this is usually of order 5–10 eV. It would be interesting to investigate if such an energy loss can be detected in experiments on the characteristic energy losses of electrons in thin foils. Since the relative intensity of the scattered primaries by this effect is of order $10^{-2}-10^{-3}/E_p$, primary energies E_p of order 10 (i.e., 272 eV) are preferable, if experiments with these energies are now feasible.

ACKNOWLEDGMENTS

I wish to thank Dr. E. M. Baroody for many fruitful discussions, and for showing me his unpublished work, Dr. A. C. Beer for suggestions on the manuscript, and Battelle Memorial Institute for financial support.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

PHYSICAL REVIEW VOLUME 134, NUMBER 6A 15 JUNE 1964

Scattering of Spin Waves by Magnetic Defects*

JOSEPH CALLAWAY AND ROBERT BOYD
Department of Physics, University of California, Riverside, California
(Received 27 January 1964)

A previous calculation of the scattering amplitude for the scattering of spin waves by magnetic defects in a simple cubic lattice is simplified and extended to body-centered and face-centered cubic lattices. Expressions are given for the mean free path, and the thermal resistivity due to defect scattering is calculated by a method which takes some account of spin-wave interactions.

I. SCATTERING THEORY

IN a previous study[1] (which will be referred to as I), the general theory of the scattering of spin waves by magnetic defects was discussed, and results were obtained for the cross section and mean free path in simple cubic lattices. In the present note, the cumbersome procedures of that calculation are both simplified and generalized, and the results are extended to include body-centered cubic and face-centered cubic lattices. This is of particular interest since several of the recently discovered ferromagnetic insulators are face-centered cubic. The thermal resistivity produced by defect scattering is calculated by a method which includes some of the effects of spin-wave–spin-wave interactions.

The theory of the scattering of spin waves by magnetic defects is an application of the general theory of the scattering of excitations by imperfections in solids. An account of this theory is being published elsewhere.[2] Therefore, the present discussion will contain only as much of the theory as is required to make the application of those results to a spin-wave system intelligible.

We consider a spin system described by a simple Heisenberg exchange Hamiltonian with nearest-neighbor interaction only:

$$H=\sum_{i,\Delta} J(\mathbf{R}_i, \mathbf{R}_i+\boldsymbol{\Delta})\mathbf{S}(\mathbf{R}_i)\cdot\mathbf{S}(\mathbf{R}_i+\boldsymbol{\Delta}). \quad (1)$$

Neither external fields nor dipole-dipole interactions are included. The system is said to contain a magnetic defect if at some site $\mathbf{R}_0$ (which we choose as the origin), there is an atom whose spin S' is coupled to its neighbors by an exchange integral J'. For the remainder of the atoms, these quantities are S and J, respectively. The quantity $\boldsymbol{\Delta}$ is a vector connecting a lattice site with one of its nearest neighbors.

The excited states of the system which contain a single spin deviation can be described by a set of functions $\phi(\mathbf{R})$ such that $|\phi(\mathbf{R})|^2$ gives the probability of finding the spin deviation on site $\mathbf{R}$. It was shown in Ref. 3 that $\phi(\mathbf{R})$ satisfies the equation

$$E\phi(\mathbf{R}_l)=2\sum_\Delta J(\mathbf{R}_l,\mathbf{R}_l+\boldsymbol{\Delta})[S(\mathbf{R}_l+\boldsymbol{\Delta})\phi(\mathbf{R}_l) -\{S(\mathbf{R}_l)S(\mathbf{R}_l+\boldsymbol{\Delta})\}^{1/2}\phi(\mathbf{R}_l+\boldsymbol{\Delta})]. \quad (2)$$

In this equation $S(\mathbf{R}_l)$ is the "magnitude" of the spin on the site $\mathbf{R}_l$ in that the eigenvalue of $S^2(\mathbf{R}_l)$ in the ground state is $S(\mathbf{R}_l)[S(\mathbf{R}_l)+1]$. Equation (2) can be written as a Schrödinger equation in the form

$$\sum_l (m|E-H_0|l)\phi(\mathbf{R}_l)=\sum_l (m|V|l)\phi(\mathbf{R}_l), \quad (3)$$

in which

$$(m|E-H_0|l)=(E-2JSz)\delta_{l,m}+2JS\delta_{l-m,\Delta} \quad (4)$$

and

$$(m|V|l)=\delta_{lm}[2Sz(J'-J)\delta_{l,0}+2(J'S'-JS)\delta_{l,\Delta}] -2[J'(S'S)^{1/2}-JS]\delta_{l-m,\Delta}(\delta_{l,0}+\delta_{m,0}). \quad (5)$$

Each atom is assumed to have z nearest neighbors. The subscript $l-m$ refers to $\mathbf{R}_l-\mathbf{R}_m$.

The solutions of this equation for energies, E, within the continuous spectrum of H_0 have been shown to have the following asymptotic form:

$$\phi(\mathbf{R}_l)=\frac{\Omega^{1/2}}{(2\pi)^{3/2}}\left[e^{i\mathbf{k}\cdot\mathbf{R}_l}+f\frac{e^{ikR_l}}{R_l}\right], \quad (6)$$

provided the energy is low enough so that the energy wave vector relation (energy band) for spin waves is spherical:

$$E=\gamma k^2, \quad (7)$$

$$\gamma=2JSa^2, \quad (8)$$

in which a is the cubic lattice constant. The quantity f may be interpreted as the scattering amplitude for spin waves. It may be expressed as a sum of scattering amplitudes for partial waves transforming according to one of the irreducible representations of the crystal point group:

$$f=\sum_\beta f_\beta, \quad (9)$$

in which β denotes an irreducible representation. The partial-wave scattering amplitudes are

$$f_\beta=-\frac{2\pi^2}{\gamma D_\beta}\sum_{pnm} V_{\beta,pn}P_{\beta,nm}\sum_\nu C_{\beta,\nu}^{(0)*}(k',R_p) \times C_{\beta,\nu}^{(0)}(k_0,R_m). \quad (10)$$

* Supported by the U. S. Air Force Office of Scientific Research.

[1] J. Callaway, Phys. Rev. 132, 2003 (1963). References to previous studies of the spin-wave–defect interaction are given there.

[2] J. Callaway, J. Math. Phys. 5, 784 (1964).

[3] T. Wolfram and J. Callaway, Phys. Rev. 130, 2207 (1963).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In this equation, the functions $C^{(0)}(k,R)$ are symmetrized linear combinations of plane waves $(\Omega/8\pi^3)^{1/2}\times\exp(i\mathbf{k}\cdot\mathbf{R})$ transforming according to the νth row of the βth irreducible representation. These functions are characterized by wave vectors $\mathbf{k}_0$ and $\mathbf{k}'$, in which $\mathbf{k}_0$ is the wave vector of the incident wave and $\mathbf{k}'$ is a vector of magnitude k_0 parallel to $\mathbf{R}_l$ [see Eq. (6)]. We can write

$$C_{\beta,\nu}^{(0)}(k,R_m)=\frac{\Omega^{1/2}}{(2\pi)^{3/2}}\sum_m{}' U(\beta_\nu,\mathbf{R}_m)e^{i\mathbf{k}\cdot\mathbf{R}_m}, \quad (11)$$

in which the quantities $U(\beta_\nu,R_m)$ are the matrix elements of a unitary transformation. The prime on the summation in Eq. (11) indicates that the vectors R_m which are included are those which can be found from any one of them by rotation with all the operators of the point group. (Only distinct vectors are considered.)

The quantities $V_{\beta,pn}$ are given by

$$V_{\beta,pn}=\sum_{p,n}{}' U(\beta_\nu,\mathbf{R}_p)(p|V|n)U^\dagger(\mathbf{R}_n,\beta_\nu), \quad (12)$$

in which $U^\dagger$ is the adjoint of the matrix U. $V_{\beta,pn}$ is independent of the row of the representation ν. The remaining quantities are found as follows. We consider the Green's function, $\mathcal{G}(\mathbf{R}_l-\mathbf{R}_n)$, which satisfies

$$\sum_l(m|E-H_0|l)\mathcal{G}(\mathbf{R}_l-\mathbf{R}_n)=\delta_{mn}, \quad (13)$$

and define

$$\Gamma_{\beta,ln}=\sum_{ln}{}' U(\beta_\nu,\mathbf{R}_l)\mathcal{G}(\mathbf{R}_l-\mathbf{R}_n)U^\dagger(\mathbf{R}_n,\beta_\nu). \quad (14)$$

Then, considering only those values of the indices l, m, such that $V_{\beta,lm}$ is not zero, we construct the matrix

$$Q_{\beta,lm}=\delta_{lm}-\sum_n\Gamma_{\beta,ln}V_{\beta,nm}. \quad (15)$$

The summation in (10) and (15) includes one term for each different symmetrized combination of plane waves. The quantities P_β and D_β occurring in Eq. (9) may now be defined by writing the matrix inverse to Q as

$$[Q^{-1}]_{\beta,lm}=D_\beta^{-1}P_{\beta,lm}, \quad (16)$$

with

$$D_\beta=\det(Q_\beta)=\det[I-\Gamma_\beta V_\beta]. \quad (17)$$

The quantity D_β determines the locations of possible scattering resonances or localized modes. Let E' (in general, a complex number) be a solution of

$$D_\beta(E')=0. \quad (18)$$

Put $E'=E_R-i\Gamma/2$. Then, if $\Gamma>0$, E_R is the energy of a scattering resonance, and Γ is the width of the resonance. If $\Gamma=0$ (which occurs only outside the continuum of eigenstates of H_0), we have a localized mode.

In the case of spin waves, the perturbation V extends only to nearest neighbors of the defect. The irreducible representations which occur in the sum over β in Eq. (7) have been determined for cubic lattices[2,3]: They are Γ_1, Γ_{15}, and Γ_{12} for a simple cubic lattice; Γ_1, Γ_{15}, $\Gamma_{25'}$, and $\Gamma_{2'}$ for a body-centered cubic lattice, and Γ_1, Γ_{15}, $\Gamma_{25'}$, Γ_{12}, and Γ_{25} for face-centered cubic.[4] The k dependence of the partial-wave scattering amplitude f_β at low energies can be studied by expanding the symmetrized combinations of plane waves $C_{\beta,\nu}^{(0)}(k,R)$ in powers of k. One finds that

$$C_{\beta,\nu}^{(0)}(k,R)\propto(kR)^lK_{\beta,\nu,l}(\theta,\phi), \quad (19)$$

in which $K_{\beta,\nu,l}$ is a Kubic harmonic belonging to the νth row of the βth irreducible representation, and is a linear combination of spherical harmonics of order l.[5] For Γ_1, $l=0$; for Γ_{15}, $l=1$; for Γ_{12} and $\Gamma_{25'}$, $l=2$, and for $\Gamma_{2'}$ and Γ_{25}, $l=3$. We conclude that (since $|\mathbf{k}_0|=|\mathbf{k}'|$)

$$f_\beta\propto k^{2l}, \quad (20)$$

unless there are cancellations among the elements V, P which raise the power of k in (20). We do not expand D_β since this contains the possible resonances.

For spin waves we find that the p-wave amplitude (Γ_{15}) is proportional to k^2, but cancellations do occur for the s-wave amplitude (Γ_1), which then turns out also to be proportional to k^2. Other (d,f) partial waves give terms proportional to k^4 and k^6. Hence, the total amplitude is proportional to k^2 for small k.

We will now determine the scattering amplitude to this order for simple cubic, body-centered cubic, and face-centered cubic lattices. The p-wave portion i simplest because only one term appears in the summation over p, m, and n in Eq. (8). From this fact x follows, with the use of Eq. (16), that the only satrim element of P is unity. Hence Eq. (10) has the for it

$$f_p=(-2\pi^2/\gamma D_p)V_{p,11}\sum_\nu C_{p,\nu}^{(0)*}(k',R_1)\times C_{p,\nu}^{(0)}(k_0,R_1). \quad (21)$$

The rows of the triply degenerate Γ_{15} representation contain functions transforming as x, y, or z. In each of the three lattices we have in the small k limit

$$C_{p_x}=[\Omega^{1/2}/(2\pi)^{3/2}]\sqrt{2}ik_xa, \quad (22)$$

etc., and

$$V_{p,11}=-2JS[1-(J'S'/JS)]. \quad (23)$$

Then

$$f_p=(\Omega/2\pi D_pa^2)[1-(J'S'/JS)]k^2a^2\cos\theta, \quad (24)$$

in which θ is the angle between k and k'.

It is now necessary to express the determinant D_p in terms of the Green's functions. If (ξ,η,ζ) are the rectangular components of $(\mathbf{R}_l-\mathbf{R}_n)/a$, we have for

[4] Notation for the irreducible representations of the point group is in accord with L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).

[5] F. C. Von der Lage and H. A. Bethe, Phys. Rev. 71, 612 (1947).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

the three lattices:

Simple cubic,

$$D_p=1-\tfrac{1}{2}[1-(J'S'/JS)][\mathcal{G}(0)-\mathcal{G}(2,0,0)]; \quad (25)$$

body-centered cubic,

$$D_p=1-\tfrac{1}{2}[1-(J'S'/JS)] \times[\mathcal{G}(0)+\mathcal{G}(1,0,0)-\mathcal{G}(1,1,0)-\mathcal{G}(1,1,1)]; \quad (26)$$

face-centered cubic;

$$D_p=1-\tfrac{1}{2}[1-(J'S'/JS)][\mathcal{G}(0)+2\mathcal{G}(\tfrac{1}{2},\tfrac{1}{2},0) -2\mathcal{G}(1,\tfrac{1}{2},\tfrac{1}{2})-\mathcal{G}(1,1,0)]. \quad (27)$$

We infer from these results that for the same value of the lattice constant a, the scattering amplitude depends on the lattice structure partly through the cell volume Ω which has the value a^3, $a^3/2$, and $a^3/4$ in the simple body-centered, and face-centered cubic lattices, respectively. In addition, more complicated dependence on structure is contained in the determinants, D_p.

We now turn to a consideration of the s-wave amplitude. This calculation is quite tedious and will not be given in detail. For each lattice, the relevant quantities V, P are 2×2 matrices. The symmetrized linear combinations of plane waves are given by

$$C_s^{(0)}(k,0)=\Omega^{1/2}(2\pi)^{-3/2}, \quad (28a)$$

$$C_s(k,R_1)=(\Omega/z)^{1/2}(2\pi)^{-3/2}\textstyle\sum_\Delta \exp(i\mathbf{k}\cdot\boldsymbol{\Delta}). \quad (28b)$$

Equation (28b) can be expressed in terms of the spin-wave energy $E(k)$, since

$$E(k)=2JS[z-\textstyle\sum_\Delta \exp(i\mathbf{k}\cdot\boldsymbol{\Delta})].$$

Hence,

$$C_s^{(0)}(k,R_1)=(\Omega/z)^{1/2}(2\pi)^{-3/2}(z-E/2JS). \quad (29)$$

Equation (29) enables us to eliminate the functions $C_s^{(0)}$ from the s-wave amplitude. In addition, it is possible to express all the combinations of Green's functions which enter into D_s and P_s in terms of $\mathcal{G}(0)$, with the aid of certain identities which were derived in Ref. 3. Since $\mathcal{G}(0)$ can be expressed in terms of the density of states $G(E)$,

$$\mathcal{G}(0)=\frac{\Omega}{(2\pi)^3}\int\frac{d^3q}{E-E(q)}=P\int\frac{G(E')dE'}{E-E'}-i\pi G(E) \equiv -g_0/4JS, \quad (30)$$

this means that the s-wave amplitude can be determined from the $G(E)$ without explicit reference to $E(k)$. The relevant identities are

$$z^{-1}\textstyle\sum_i \mathcal{G}(\boldsymbol{\Delta}_i)=[1-(E/2JSz)]\mathcal{G}(0)+(1/2JSz) \equiv -g_1/4JS, \quad (31)$$

$$\textstyle\sum_j \mathcal{G}(\boldsymbol{\Delta}_i-\boldsymbol{\Delta}_j)=[1-(E/2JSz)]\sum_i \mathcal{G}(\boldsymbol{\Delta}_i) =-g_2/4JS. \quad (32)$$

In these equations, $\boldsymbol{\Delta}_i$ and $\boldsymbol{\Delta}_j$ indicate sites which are nearest neighbors of the defect site (the origin). Only the particular combinations of Green's functions appearing in (30), (31), and (32) are involved in the s-wave amplitude.

In addition, we have

$$V_s=-4JS\begin{pmatrix}\epsilon & \eta\sqrt{z}\\ \eta\sqrt{z} & \rho\end{pmatrix}, \quad (33)$$

$$P_s=\begin{pmatrix}1-g_2\rho-z\eta g_1 & z^{1/2}(g_0\eta+g_1\rho)\\ z^{1/2}(g_1\epsilon+g_2\eta) & 1-g_0\epsilon-z\eta g_1\end{pmatrix}, \quad (34)$$

in which

$$\epsilon=\tfrac{1}{2}z(1-J'/J), \quad (35a)$$

$$\eta=\tfrac{1}{2}[(J'/J)(S'/S)^{1/2}-1], \quad (35b)$$

$$\rho=\tfrac{1}{2}(1-J'S'/JS). \quad (35c)$$

After a rather tedious calculation, we obtain for the s-wave portion of the scattering amplitude to lowest order in k

$$f_s=\frac{-\Omega}{2\pi D_s a^2}\frac{E}{4JS}\left(\frac{J'}{J}\right)\left(1-\frac{S'}{S}\right) =\frac{-\Omega}{4\pi D_s a^2}\left(\frac{J'}{J}\right)\left(1-\frac{S'}{S}\right)k^2a^2, \quad (36)$$

in which

$$D_s=1+\frac{z}{2}(g_1-g_0)\left(1-\frac{J'}{J}\right)+\frac{Eg_1}{4JS}\left(1-\frac{J'S'}{JS}\right). \quad (37)$$

It will be observed that f_s goes to zero as k^2 at low energies. This is the result of cancellation among some of the terms and of the Green's function identities (32). The expression (37) for D_s is exact. We do not expand D_s (or D_p) since they contain the possible low-energy resonances.

The total amplitude to order k^2 is the sum of (36) and (24):

$$f=f_s+f_p=\frac{-\Omega}{4\pi a^2}\left[\frac{1}{D_s}\left(\frac{J'}{J}\right)\left(1-\frac{S'}{S}\right) -\frac{2}{D_p}\left(1-\frac{J'S'}{JS}\right)\cos d\right]k^2a^2. \quad (38)$$

It will be observed that the amplitude in this order depends on the lattice structure in an obvious manner through the cell volume. There is, however, an equally important structure dependence incorporated in the function D_s and D_p. The higher order (in k) terms in the amplitude also exhibit a significant dependence on the lattice structure since different representations appear. In general, it does not appear to be possible to express the scattering amplitude solely in terms of the

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

पुस्तकालय 111678

angle between the incoming and outgoing wave vectors when higher order terms are included.

When the scattering amplitude is considered only to lowest order in $\hbar$, so that it depends only on the angle θ a simple calculation of the mean free path l_D, for scattering of a spin wave by a defect is possible. This has been shown to be determined by the momentum transfer cross section σ_n which is given by

$$\sigma_n=\int |f(\theta)|^2(1-\cos\theta)d\Omega. \tag{39}$$

Then[6]

$$l_D^{-1}=N_D\sigma_n, \tag{40}$$

where N_D is the concentration of defects. We obtain

$$\sigma_n=\frac{\Omega^2k^4}{4\pi}\left\{\frac{1}{|D_s|^2}\left(\frac{J'}{J}\right)^2\left(1-\frac{S'}{S}\right)^2 +\frac{4}{3}\left(1-\frac{J'S'}{JS}\right)\left[\frac{1}{|D_p|^2}\left(1-\frac{J'S'}{JS}\right) +\frac{1}{2}\left(\frac{J'}{J}\right)\left(1-\frac{S'}{S}\right)\left(\frac{1}{D_s^*D_p}+\frac{1}{D_sD_p^*}\right)\right]\right\}. \tag{41}$$

If we suppose that the denominators D_s and D_p can be set equal to unity, thereby neglecting the possibility of resonant scattering, we obtain

$$\sigma_n=\frac{\Omega^2k^4}{4\pi}\left[\left(\frac{J'}{J}\right)^2\left(1-\frac{S'}{S}\right)^2 +\frac{4}{3}\left(1-\frac{J'S'}{JS}\right)\left(1+\frac{J'}{J}-\frac{2J'S'}{JS}\right)\right]. \tag{42}$$

II. THERMAL CONDUCTIVITY

In this section we consider the thermal conductivity of a system of spin waves in interaction with each other and containing N_D magnetic defects per unit volume. We hope to consider spin-wave–phonon interactions in subsequent work. Unless explicitly stated, the system will be assumed to be of infinite extent.

We propose here to calculate the thermal conductivity of spin waves by the same techniques which have proved to be reasonably successful in application to lattice thermal conductivity.[7] We will consider each scattering process to be described by a relaxation time, but explicit account will be taken of the conservation of crystal momentum by normal spin-wave–spin-wave scattering processes.

The distinction between normal and umklapp processes in spin-wave interactions is just as vital as it is in the case of phonon-phonon interactions. Since normal processes conserve the total pseudomomentum of the spin-wave system, they cannot, if acting alone, produce a thermal resistance. This statement is essentially independent of the details of the dispersion relation for the magnons. It is true that the interactions between magnons do not conserve the heat flow at the microscopic level (as is the case for phonons if a linear dispersion law is assumed). A statistical argument shows, however, that a thermal resistance will not appear. An explicit demonstration is given in the Appendix.

Nevertheless, it is not permissible to neglect normal processes in studying the thermal conductivity. Their contribution may be pictured qualitatively as that of converting some of the low-frequency magnons into high-frequency magnons which may easily be scattered by defects. We will show that in the limit of strong normal processes, the thermal conductivity becomes limited by the defect scattering in a fashion similar to that obtained by Ziman[8] for lattice thermal conductivity.

The importance of distinguishing normal and umklapp processes has been stressed here because it has been ignored in other calculations of the spin-wave thermal conductivity.[9] We contend that such work may have yielded an erroneous thermal resistance.

We begin by considering the Boltzmann equation in the form

$$\left(\frac{\partial N}{\partial t}\right)_c-\mathbf{V}_\mathbf{k}\cdot\nabla T\frac{dN}{dT}=0, \tag{43}$$

in which ∇T is the temperature gradient, N is the distribution function, and $\mathbf{V}_\mathbf{k}$ is the group velocity of the spin waves

$$\mathbf{V}_\mathbf{k}=\hbar^{-1}\nabla_\mathbf{k}E(\mathbf{k}). \tag{44}$$

The first assumption of the present approach is that the collision term $(\partial N/\partial t)_c$ may be approximated as follows:

$$\left(\frac{\partial N}{\partial t}\right)_c=\frac{N(\boldsymbol{\lambda})-N}{\tau_n}+\frac{N_0-N}{\tau_u}, \tag{45}$$

in which $N(\boldsymbol{\lambda})$ is a displaced Bose distribution

$$N(\boldsymbol{\lambda})=\{\exp[(E-\boldsymbol{\lambda}\cdot\mathbf{k})/KT]-1\}^{-1}. \tag{46}$$

N_0 is the usual Bose function; τ_n is the relaxation time for a single mode via normal processes, and τ_u is the relaxation time for all those processes which do not conserve the crystal momentum.[10] We assume that $\boldsymbol{\lambda}$ is small so that $N(\boldsymbol{\lambda})$ may be expanded, and only first-

[6] J. M. Ziman, *Electrons and Phonons* (Oxford University Press, Oxford, 1960).

[7] J. Callaway, Phys. Rev. **113**, 1046 (1959); J. Callaway and H. C. Von Baeyer, *ibid.* **120**, 1149 (1960).

[8] J. M. Ziman, Can. J. Phys. **34**, 1256 (1956).

[9] A. Quattropani, Phys. Kondens. Materie **1**, 125 (1963).

[10] R. E. Nettleton, Phys. Rev. **132**, 2032 (1963), has partially justified this approximation for lattice thermal conductivity. In applying Nettleton's discussion to the present calculation, one should note that in spite of appearances, an explicit expression for the normal process relaxation time, τ_N, is not required to obtain our essential results, Eqs. (60) and (67). The quantities which are required are τ_D and τ_c.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

order terms in $\mathfrak{z}$ are retained. Put

$$\mathfrak{z}=-S\nabla T/T, \quad (s \text{ is a constant}) \tag{47a}$$

$$N-N_0=n_1=-\nu\mathbf{V}_k\cdot\nabla T\left(\frac{E}{KT}\right)^2\frac{e^{E/KT}}{(e^{E/KT}-1)^2}, \tag{47b}$$

$$\tau_c^{-1}=\tau_u^{-1}+\tau_n^{-1}. \tag{47c}$$

(These equations are definitions of s, ν, and τ_c.) We also require the group velocity. We continue to use a quadratic dispersion law

$$E=\gamma k^2+\epsilon, \tag{48}$$

so that

$$\mathbf{V}_k=2(\gamma/\hbar)\mathbf{k}. \tag{49}$$

The constant term ϵ in (48) does not play an important role in the mechanics of the thermal conductivity calculation and has been ignored in computing scattering amplitudes, but it is important in preventing the quantity $E-\mathfrak{z}\cdot\mathbf{k}$ in (46) from becoming negative for low energies. It is therefore important to realize that a nonzero value of ϵ is always obtained, since the material is magnetized. One has, approximately,[11]

$$\epsilon=g\beta[(4\pi/3)M+H_{\text{ext}}], \tag{50}$$

where M is the magnetization and H_{ext} is a possible external field.

Equations (45)–(49) are substituted in the Boltzmann equation, which then gives a relation between ν and s. This is

$$\nu=\tau_c[1+\hbar s/(2\gamma\tau_N E)]. \tag{51}$$

The thermal conductivity of the system is

$$\begin{aligned}\kappa&=[1/(2\pi)^3]\int \mathbf{V}_k^2\nu(k)\cos^2\theta C_{sw}(k)d^3k,\\ &=(2\gamma^2/3\pi^2\hbar)\int k^4\nu(k)C_{sw}(k)dk,\\ &=(2\gamma^2/3\pi^2\hbar^2)\int k^4\tau_c[1+S\hbar/(2\gamma\tau_n E)]C_{sw}(k)dk.\end{aligned} \tag{52}$$

We have assumed cubic symmetry and also employed (49). $C_{sw}(k)$ is the contribution to the specific heat from a mode of wave vector k. Evidently another relation between ν and s is required. This is obtained from the requirement that the normal processes do not change the total wave vector of the system:

$$\int\left(\frac{\partial N}{\partial t}\right)_N\mathbf{k}d^3k=\int\frac{N(\mathfrak{z})-N}{\tau_N}\mathbf{k}d^3k=0. \tag{53}$$

This condition leads to the relation

$$\int\frac{\mathbf{k}}{\tau_N}(\mathbf{k}\cdot\nabla T)\left[s-\frac{2\gamma\nu E}{\hbar}\right]\frac{e^{E/KT}}{(e^{E/KT}-1)^2}d^3k=0. \tag{54}$$

Only the component of $\mathbf{k}$ parallel to the temperature gradient survives on integration over solid angles. We substitute (51) and solve for s

$$s=\frac{2\gamma}{\hbar}\int\frac{\tau_c}{\tau_N}\frac{Ek^4e^{E/KT}}{(e^{E/KT}-1)^2}dk\Bigg/\left[\int\frac{1}{\tau_N}\left(1-\frac{\tau_c}{\tau_N}\right)\frac{k^4e^{E/KT}}{(e^{E/KT}-1)^2}dk\right]. \tag{55}$$

We combine (55) and (53), substitute for the spin-wave specific heat

$$C_{sw}=K\left(\frac{E}{KT}\right)^2\frac{e^{E/KT}}{(e^{E/KT}-1)^2},$$

introduce the change of variables

$$x=\gamma k^2/KT, \quad \delta=\epsilon/KT, \tag{56}$$

replace the relaxation time by mean free paths,

$$l_c=V_k\tau_c,$$

etc., and obtain our final expression for the thermal conductivity:

$$\kappa=\frac{K(KT)^2}{6\pi^2\gamma\hbar}\left\{\int_0^{x_m}x(x+\delta)^2l_c\frac{e^{x+\delta}}{(e^{x+\delta}-1)^2}dx+\left[\int_0^{x_m}x^{3/2}\frac{l_c}{l_N}\frac{(x+\delta)e^{x+\delta}}{(e^{x+\delta}-1)^2}dx\right]^2\Bigg/\left[\int_0^{x_m}\frac{l_c}{l_nl_u}\frac{x^2e^{x+\delta}}{(e^{x+\delta}-1)^2}dx\right]\right\}. \tag{57}$$

The upper limit x_m for these integrals is determined by an argument analogous to that establishing the Debye temperature for a crystal lattice and is given by $x_m=(\gamma/KT)(6\pi^2/\Omega)^{2/3}$. We will consider here only temperatures low enough so that the upper limit can be made infinite.

The dependence of the thermal conductivity on an external magnetic field can be obtained from (57). A magnetic field decreases the thermal conductivity by reducing the probability of excitation of spin waves. It is evident from (57) that in the limit of large δ (low temperatures or strong fields), the thermal conductivity must decrease as $\delta^ne^{-\delta}$; here the exponent n will depend on the scattering mechanism. Further discussion of the field dependence of the conductivity will be reserved for subsequent work.

[11] T. Holstein and Primakoff, Phys. Rev. 58, 1098 (1940); S. H. Charap and E. Boyd (to be published). This dispersion relation is only approximate since in the presence of dipole-dipole couplings, the spin-wave energies are anisotropic.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

There is a great similarity between Eq. (57) and the corresponding expressions which are obtained in the theory of lattice thermal conductivity. In particular, we observe that if the relaxation time for processes which do not conserve the wave vector becomes very long, the thermal conductivity of the spin-wave system increases without limit. If only normal processes were present, there would be no thermal resistance.

It follows that in the limit in which the mean free path for normal processes is very short compared to that for processes which do not conserve the wave vector, the thermal conductivity will depend sensitively on the latter processes. This is the Ziman limit. To obtain the thermal resistance in this limit, we drop the first term of (57) and neglect the contribution of l_u to l_c. Then

$$W=\frac{1}{\kappa}=\frac{6\pi^2\gamma\hbar}{K(KT)^2}\int\frac{x^2}{l_u}\frac{e^{x+\delta}}{(e^{x+\delta}-1)^2}dx\Big/\left[\int x^{3/2}\frac{(x+\delta)e^{x+\delta}}{(e^{x+\delta}-1)^2}dx\right]^2. \quad (58)$$

One interesting consequence of (58) is that the thermal resistance in the Ziman limit is the sum of additive contributions from each possible type of "U" process. In particular we consider defect scattering, for which the mean free path is given by Eqs. (40) and (41). We will not consider here the effects of possible scattering resonances which could increase the thermal resistance substantially. We abbreviate the mean free path due to defect scattering as

$$l_D^{-1}=N_DAk^4, \quad (59a)$$

in which

$$A=\frac{\Omega^2}{4\pi^2}\left[\left(\frac{J'}{J}\right)^2\left(1-\frac{S'}{S}\right)^2+\frac{4}{3}\left(1+\frac{J'}{J}-2\frac{J'S'}{JS}\right)\right]. \quad (59b)$$

We can evaluate the thermal resistance due to defects approximately by neglecting δ in (58). We get

$$W_D=\frac{6\pi^2\hbar N_DA}{K\gamma}\int_0^\infty\frac{x^4e^x}{(e^x-1)^2}dx\Big/\left[\int_0^\infty\frac{x^{5/2}e^x}{(e^x-1)^2}dx\right]^2$$
$$=\frac{6\pi^2N_DA\hbar}{K\gamma}\frac{\Gamma(5)\zeta(4)}{[\Gamma(7/2)\zeta(5/2)]^2}=38.6\frac{N_DA\hbar}{JSKa^2}. \quad (60)$$

In the last step of (60), we have substituted Eq. (8).

The thermal resistance due to the defects is independent of temperature. This expression is valid in the limit in which defect scattering is weak compared to normal magnon-magnon scattering. Our treatment has included the effects of magnon-magnon scattering on the distribution function, and thereby has taken account of these interactions implicitly. Our formula, Eq. (60), does not depend explicitly on the details of the magnon-magnon interaction. It does differ somewhat from the expression for the defect resistivity if these interactions are neglected. The situation in which defect scattering is weak compared to boundary scattering and normal processes are ignored was examined in I. In that case a thermal resistance W_D' due to defects was obtained which is independent of temperature, and differs from (60) only in respect to the numerical constant. That result is[12]

$$W_D'=20\pi^2\frac{N_DA\hbar}{K\gamma}\frac{\zeta(5)}{[\zeta(3)]^2}=70.8\frac{N_DA\hbar}{JSKa^2}. \quad (61)$$

It will be noted that the thermal resistance due to defects given in (61) is larger by a factor of almost 2 than that of Eq. (60). This situation should be contrasted with that which obtains in the theory of lattice thermal conductivity where the contribution to the thermal resistance from defects is, in the Ziman limit, larger by a factor of 25 than when umklapp scattering dominates.

Next, we will discuss the thermal conductivity in the opposite limit in which the defect scattering is strong compared to magnon-magnon scattering. This requires a more detailed knowledge of the combined mean free path l_c, including normal processes. The interaction of two spin waves can produce bound states,[13] and probably scattering resonances. A complete theory of the scattering of two spin waves has not yet been given, but for an introduction, we can make use of a calculation reported by Dyson.[14] Dyson showed that, to the extent that resonances and such may be neglected, the mean free path for a spin wave of wave vector k due to spin-wave–spin-wave interactions, may be written as

$$l=BT^{5/2}k^2. \quad (62)$$

This expression is valid at low temperatures. The quantity B is a rather complicated function of J and S, and is given by

$$B=\frac{3\zeta(3/2)a}{z\omega S^2}\left(\frac{3K}{2\pi JSz\mu}\right)^{5/2}, \quad (63)$$

in which

$$\omega=(1, 3\times2^{-5/3}, 2^{-1/3}),$$
$$\mu=(1, 3\times2^{-4/3}, 2^{1/3}), \quad (64)$$

for the simple cubic, body-centered cubic, and face-centered cubic lattices, respectively.

With the use of (62) and (59), we have a combined mean free path including defect scattering and magnon-magnon scattering

$$l_c^{-1}=BT^{5/2}k^2+N_DAk^4. \quad (65)$$

[12] A factor of $(K\zeta(3))^{-1}$ was omitted in the statement of this equation in Ref. 1.

[13] M. Wortis, Phys. Rev. 132, 85 (1963); J. Hanus, Phys. Rev. Letters 11, 336 (1963).

[14] F. J. Dyson, Phys. Rev. 102, 1217 (1956).

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

In the limit $l_c \ll 1$ which corresponds to defect scattering strong compared to magnon-magnon scattering, we can drop all but the first term of (57). This follows since if we consider both l_c and l_u to be proportional to some small constant α, the first term of (57) is proportional to α, whereas the second term is proportional to α^2. We then substitute (65) into (57), let $\delta \to 0$, and obtain

$$\kappa = \frac{K\gamma}{6\pi^2\hbar N_D A}\int_0^{x_m} \frac{x^2}{x+y}\frac{e^x}{(e^x-1)^2}dx, \qquad (66)$$

in which $y = B\gamma T^{3/2}/N_D A K$.

We consider this integral at temperatures low enough so that the upper limit may be made infinite. Even so, the integral needs to be approximated if we are to avoid numerical computation. We follow the procedure used under somewhat similar circumstances in I, in which we replace $x^2e^x/(e^x-1)^2$ by 1 for $x<1$, and neglect y for $x>1$. This is valid only when $y \ll 1$, but if this condition is not satisfied, it is not legitimate to neglect the second term in (57). Hence

$$\kappa = \frac{K\gamma}{6\pi^2\hbar N_D A}\left[\int_0^1 \frac{dx}{x+y} + \int_1^\infty \frac{xe^x}{(e^x-1)^2}dx\right]$$

$$= \frac{K\gamma}{6\pi^2\hbar N_D A}\left[\ln\left(1+\frac{1}{y}\right) + \sum_{n=1}^{\infty} e^{-n}\left(1+\frac{1}{n}\right)\right]$$

$$= \frac{KJSa^2}{3\pi^2 N_D A\hbar}\left[\ln\left(1+\frac{N_D A K}{B\gamma T^{3/2}}\right) + 1.0407\right]. \qquad (67)$$

Evidently, in this limit, the temperature dependence of the thermal conductivity is weak.

APPENDIX

In this appendix, we give an explicit demonstration that for magnons whose energy is related to the wave vector by Eq. (48), normal scattering processes (N. P.) cannot, by themselves, produce a thermal resistance. We recall that a normal process is one in which the total wave vector of the excitations involved is conserved. It is to be distinguished from an umklapp process, for instance, in which the wave vector changes by a reciprocal lattice vector.

It is fairly easy to see intuitively that if only N. P. operate, the crystal relaxes to equilibrium subject to the condition that the total crystal wave vector be nonzero and constant $\sum_k \mathbf{k}n_\mathbf{k} = \mathbf{K}$ (in which $n_\mathbf{k}$ is the number of particles in a state of wave vector $\mathbf{k}$). There must be a flow of heat in the direction of $\mathbf{K}$.

The simplest mathematical justification of this intuition can be found by returning to the derivation of the Fermi-Dirac (F. D.) and Bose-Einstein (B. E.) distribution functions for an ideal gas using the microcanonical ensemble. A system of weakly interacting excitations is considered.

To obtain the distribution function n_k we maximize the entropy, σ, subject to the constraints imposed by conservation laws.[15] We do this by multiplying those quantities $g_j(n_\mathbf{k})$ which must be conserved by Lagrange undetermined multipliers λ_j. Then we add their variation with respect to $n_\mathbf{k}$ to the variation of σ and set the whole equal to zero. Thus,

$$\frac{\delta\sigma}{\delta n_\mathbf{k}} + \sum \lambda_j \frac{\delta g_j}{\delta n_\mathbf{k}} = 0. \qquad (A1)$$

Let the upper/lower signs represent B. E./F. D. statistics, respectively. We have

$$\delta\sigma/\delta n_\mathbf{k} = \ln(n_\mathbf{k}^{-1} \pm 1). \qquad (A2)$$

Let us choose the following conserved quantities:

$$\sum_\mathbf{k} n_\mathbf{k} = N; \quad \sum_\mathbf{k} E_\mathbf{k} n_\mathbf{k} = E; \quad \sum_\mathbf{k} \mathbf{k} n_\mathbf{k} = \mathbf{K}. \qquad (A3)$$

We vary these quantities, multiply by Lagrange multipliers, and substitute together with (2) in (1) to obtain

$$n_\mathbf{k} = \left[\exp\left(\frac{E_\mathbf{k} - \mu - \boldsymbol{\lambda}\cdot\mathbf{k}}{KT}\right) \mp 1\right]^{-1}, \qquad (A4)$$

where the multipliers of N, E, and $\mathbf{K}$ have been written in the conventional manner. (If the number of particles is not conserved, we should set $\mu = 0$, and this must be done for magnons and phonons.)

Clearly, since $E_\mathbf{k} = E_{-\mathbf{k}}$, modes with $\mathbf{k}$ parallel to $\boldsymbol{\lambda}$ contain more particles than modes with $\mathbf{k}$ antiparallel, so we expect a net heat current Q in the direction of $\boldsymbol{\lambda}$.

To demonstrate this we consider the expression for the heat current

$$\mathbf{Q} = \sum_k E_\mathbf{k} V_\mathbf{k} n_\mathbf{k}. \qquad (A5)$$

We will evaluate this approximately on the basis of the dispersion relation[16]

$$E_k = \gamma k^n + \epsilon. \qquad (A6)$$

In this equation the relevant values of the exponent n are 1 and 2. If $n=1$, we have a situation approximately describing low-energy phonons (also $\epsilon = 0$ in this case). The case $n=2$ characterizes spin waves, but now $\epsilon \neq 0$; a nonzero value of E is provided by the magnetization (M) plus any external field (H_{ext}), as was described in the main text [see Eq. (50)]. With the use of (6), we obtain for Q

$$\mathbf{Q} = \frac{n\gamma}{(2\pi)^3\hbar}\int \frac{(\gamma k^n + \epsilon)\mathbf{k}k^{n-2}}{\{\exp\{[\gamma k^n + \epsilon - \boldsymbol{\lambda}\cdot\mathbf{k}]/KT\} - 1\}}d^3k. \qquad (A7)$$

---

[15] K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), p. 192.

[16] The importance of including ϵ is that if it is not present, the argument of the exponential in (A4) will become negative for small k when $n=2$.

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

111678

Digitized by Arya Samaj Foundation Chennai and eGangotri

The total wave vector K is

$$\mathbf{K}=\frac{1}{(2\pi)^3}\int\frac{\mathbf{k}}{\{\exp\{[\gamma k^n+\epsilon-\boldsymbol{\lambda}\cdot\mathbf{k}]/KT\}-1\}}d^3k. \quad \text{(A8)}$$

Neither of these integrals is zero. If $n=1$, $(\epsilon=0)$ it is easy to see that

$$\mathbf{Q}=(\gamma/\hbar)\mathbf{K}. \quad \text{(A9)}$$

For $n=2$, $(\epsilon\neq 0)$ the situation is somewhat more complex. The substitution $\mathbf{q}=\mathbf{k}-\boldsymbol{\lambda}/2\gamma$ enables the integrals to be transformed to a form in which the portions which vanish on integration over angle can be readily separated. The results are

$$\mathbf{Q}=\frac{\boldsymbol{\lambda}}{4\pi^2\hbar}\left(\frac{KT}{\gamma}\right)^{3/2}\left[\frac{5KT}{3}I(x_0,\tfrac{3}{2})+\left(\epsilon+\frac{\lambda^2}{4\gamma}\right)I(x_0,\tfrac{1}{2})\right], \quad \text{(A10a)}$$

$$\mathbf{K}=\frac{\boldsymbol{\lambda}}{8\pi^2\gamma}\left(\frac{KT}{\gamma}\right)^{3/2}I(x_0,1/2), \quad \text{(A10b)}$$

in which

$$I(x_0,\nu)=\Gamma(\nu+1)\sum_{n=1}^{\infty}\frac{e^{-nx_0}}{n^{\nu+1}} \quad \text{(A11)}$$

and $KTx_0=\epsilon-\lambda^2/4$.

Equation (A10b) may be used to eliminate $\boldsymbol{\lambda}$ from the expression for $\mathbf{Q}$. In the limit in which ϵ and λ are both small, we obtain a simple result

$$\mathbf{Q}=C(KT/\hbar)\gamma\mathbf{K}, \quad \text{(A12)}$$

in which C is a numerical constant.

We have shown that in the absence of processes which do not conserve the total wave vector, there is a constant nonzero heat current in the presence of nonzero total crystal wave vector. A state with a nonvanishing $\mathbf{K}$, once established, could not be changed by collisions. Since no thermal gradient was assumed, this means that in the absence of umklapp processes the ideal no thermal resistivity. There is no conflict with the result of the analogous classical problem of the thermal conductivity of gas in a cylinder, since in that case a state of nonzero momentum is not established.

पुस्तकालय
गुरुकुल कांगड़ी विश्वविद्यालय
हरिद्वार

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

(Aii)

λ from
nd λ are

(A12)

rocesses
here is a
of non-
nvanish
nged by
ned, this
s the
wit. e
thermal
t case a

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri

Compiled
1999-2000

CC-0. In Public Domain. Gurukul Kangri Collection, Haridwar

Digitized by Arya Samaj Foundation Chennai and eGangotri